

OBITUARY NOTICE

PROFESSOR WILHELM STERN, 1871-1938

WE regret to announce the death of Professor Wilhelm Stern, an Honorary Member of the British Psychological Society. He had been associated mainly with the University of Hamburg. But he was forced to leave Germany, and went to work in the Department of Psychology at Duke University. He is best known in this country for his book, *The Psychology of Early Childhood* (first German edition, 1914), which is regarded as a standard work on the subject. His earlier work on applied psychology—mental measurement, vocational guidance, etc.—was succeeded by his main life's work on the 'science of personalistics'. This included approaches from the psychological, physiological, cultural and metaphysical aspects of 'personalism', which are embodied in a three-volume work, published between 1906 and 1924, *Person und Sache: System des kritischen Personalismus*. It was further developed as a systematic psychology in *Allgemeine Psychologie auf personalistischer Grundlage* (1935), translated and published in 1938 as *General Psychology from the Personalistic Standpoint*.

THE CO-OPERATION OF SOCIAL GROUPS. PART I

BY D. M. CARMICHAEL

From the Cambridge Psychological Laboratory

- I. *A scheme of research* (pp. 206–209).
 - (i) *Direct observation of limited communities* (pp. 207–208).
 - (ii) *Controlled experimentation* (p. 208).
 - (iii) *Experiments in constructive social thinking* (pp. 208–209).
- II. *Further experiments in constructive social thinking* (pp. 210–215).
 - (i) *Material of Case I: "Church Union in Thorness"* (pp. 210–214).
 - (ii) *Method of presentation* (p. 215).
- III. *Analysis of results of Case I* (pp. 216–231).
 - (1) *Numerical analysis* (pp. 216–218).
 - (2) *Qualitative analysis* (pp. 218–231).
 - (i) *Initial union* (pp. 219–222).
 - (ii) *Permanent union* (pp. 223–225).
 - (iii) *Impermanent union* (pp. 225–226).
 - (iv) *Harmonious union* (pp. 226–227).
 - (v) *Union with disharmony* (pp. 227–228).
 - (vi) *The fisher-folk* (pp. 228–229).
 - (vii) *The influence of a leader* (pp. 229–231).

I. A SCHEME OF RESEARCH

EXACTLY 200 years ago David Hume conceived of a science of man based not on unverifiable metaphysical speculation but on accurate observation and the use of the experimental method. In this science "we must glean up our experiments from a cautious observation of human life, and take them as they appear in the common course of the world, by men's behaviour in company, in affairs and in their pleasures. When experiments of this kind are judiciously collected and compared, we may hope to establish on them a science which will not be inferior in certainty, and will be much superior in utility, to any other of human comprehension."

Hume's high hopes have scarcely been realized. General psychology, it is true, has gone a long way towards perfecting experimental methods even more objective than those which Hume seems to have in mind, but

of the utility of much of this work the ordinary man in the street has still to be convinced. In social psychology, on the other hand, where the utility of scientific investigation is more immediately apparent, the complexity of the data and the difficulty of performing controlled experiments are such that individual interpretation and speculation still hold the field, and "even the rabble without doors may judge from the noise and clamour which they hear, that all goes not well within".

The experiments described in this paper form part of a wider scheme of research devised and inaugurated at Cambridge by Prof. F. C. Bartlett. This scheme aims ultimately at discovering, as far as possible by experimental methods, the psychological factors which govern social harmony and discord, with special reference to those conditions which make for permanent and friendly co-operation between differently organized social groups. A solution of this all-important question is being sought at least as much for its practical value as for its academic interest. The methods of approach which it is proposed to use, and the reasons for their adoption, can best be indicated by giving a brief summary of a "Preliminary Report and Suggestions" which Prof. Bartlett has already published.¹

Three interdependent lines of research are suggested, of which the third is intended to give direction to the first and second, and the first and second to validate the third.

(i) *Direct observation of limited communities*

"There is a large number of limited communities within which the friendly co-operation of different social groups is demanded", either permanently or from time to time. The history, outstanding membership and cultural background of many of these small groups is either already well known, or can easily be determined, and the interaction of the groups can therefore be at once made the subject of first-hand observation. This observation will be directed towards discovering "what are the most common *human* conditions which either maintain, or hinder, friendly co-operation".

All such observational work will, of course, "have to pay attention to certain principles of method in the collection and recording of facts", and it would be desirable to have the facts "available at some common centre in a form which renders their evaluation possible". This, it is interesting to note, is precisely the method proposed by Hume—the

¹ "The co-operation of social groups" (1938), *Occupational Psychol.* xii, 30-42.

"judicious collection" of relevant data "as they appear in the common course of the world", and their subsequent comparison and utilization.

(ii) *Controlled experimentation*

Within the type of community mentioned, it is possible to stage social experiments either with groups which are already in contact, or with groups which have been brought together 'for the occasion', and encouraged to co-operate¹. In both these instances, the experimenter could "exaggerate, or otherwise alter certain outstanding cultural or individual conditions which are suspected to be of influence, and then keep a watch upon their operation".

As it seems to be a reasonable and necessary working hypothesis to suppose that the *principles* which govern the co-operation and conflict of groups remain much the same, whatever the size of the group, the fact that observational and experimental work can be carried out only with small groups is relatively unimportant. What is important, however, is that this observational and experimental work should be properly directed. The mere amassing of data defeats its own ends, and the conducting of significant experiments becomes impossible unless some principle of selection is used which enables us, within limits, to distinguish between those factors which are relevant and those which can safely be neglected. The third method of approach, now to be described, will, it is hoped, provide principles of selection.

(iii) *Experiments in constructive social thinking*

This method aims primarily at discovering what factors are *commonly believed* to be of leading importance in the relationships of groups. In a sense which will become apparent, it is essentially an attempt to discover "how people think about social questions". Bartlett (*loc. cit.*) describes it by stating the manner of its adoption: his experiments on remembering having revealed "certain very important processes of mental construction", he thought that "it would be interesting to get to closer grips with these processes, to find how and when they occur, to distinguish between the part of them which is contributed by the individual and the part, if any, which has a universal determination". With this end in view, he "devised and applied a considerable variety of stimulus material", including certain (written) descriptions "in which situations were depicted involving more or less indeterminate relationships between two or more people and the subjects were asked to carry

¹ Cf. Lewin, K. and Lippit, R. (1938). *Sociometry*, 1, 292-300.

further the probable issue of these relationships". It was noticed that "when they did this they very often made use of facts of a genuinely cultural or group character, which could not be expressed in terms of individual behaviour or experience alone". It was therefore decided "to have instances involving direct group relationships, and to get the subjects to think constructively about them".

A number of concrete situations were therefore described, all involving group relationships not unlikely to occur within the experience of the subjects available. These situations "were presented in an incomplete form, and the subjects were asked to carry further their probable history". The "Preliminary Report and Suggestions" prints two of the descriptions used by Bartlett, and contains an analysis of the results which they yielded. It is unnecessary here to discuss these results in detail, as one of the descriptions in question has since been used more extensively, and reference to it will be made at a later stage in this paper. But Bartlett's general conclusions as regards his new method are important. They are three in number:

(1) "When an ordinary person is confronted by a concrete description of a social situation in which agreement between two groups is sought, but some discordant elements are present, he can usually give a considered opinion as to the probable upshot." The method has, for this reason, certain marked advantages over most questionnaires.

(2) "The method shows that there can be a considerable measure of agreement regarding the primary bases of social harmony and discord, and at the same time a wide diversity of view about the particular mechanism through which these fundamental conditions find expression." A further and wider application of the method will, it is hoped, sharpen and define this important practical distinction.

(3) A fairly circumscribed application of the method has already demonstrated that "ideas about the leading conditions governing co-operation and discord may vary significantly from one section to another of what is considered the same general community".

Widespread views about social co-operation must undoubtedly affect the chances of social co-operation, and therefore the opinions elicited by our third method will be to that extent of practical value. Whether the opinions are true or false, however, can be determined only by the use of the first two methods of direct observation and controlled experimentation. If, by the combined use of the three methods we are ever able to discover the factors which actually govern social co-operation, then the practical value of the whole enquiry will be immensely increased.

II. FURTHER EXPERIMENTS IN CONSTRUCTIVE SOCIAL THINKING

I shall now proceed to describe and discuss further experiments which have been carried out along the lines suggested by Bartlett. Three 'descriptions' will be referred to—one which Bartlett has already tested; one, a leading article from *The Times*, dealing with an actual attempt to secure co-operation between two sporting associations; and one of my own composition, which is given below as 'Case I'. In each instance I shall (1) reprint the material used; (2) state the reasons for its adoption and the manner of its presentation; (3) give a brief analysis of results and attempt to assess their significance. Finally, I shall endeavour to estimate the value of the work which has so far been done in relation to the wider scheme of research of which it forms a part.

(i) *Material of Case I: "Church Union in Thorness"*

In 1929, after a long series of legal, financial and administrative adjustments, and one or two minor changes in dogma, the United Free and the old Established Church of Scotland became in name and in the eyes of the state, one church. The day was indeed a memorable one when, for the first time, the representatives of each Church met as one body, and the first Moderator of this first General Assembly of the new 'Church of Scotland' rose to address its members. He spoke of the work which had been done, and of the success with which it had been crowned; of the difficulties which had been overcome; of the unselfishness and ceaseless vigilance of those to whom had been entrusted the task of framing the Act of Union; of the inspiration and guidance which had been granted them as they strove to realize their high ideal, and of the whole-hearted support which they had received from ministers of both denominations, and from eminent laymen within the churches. It was impossible, he said, to consider these things, and to look upon that vast assemblage, and not to feel convinced that, within the body of the ministry at least, the old prejudices, the old hatreds, the old jealousies had perished, and along with them all the waste, the bitterness, the pride and the self-righteousness born of two rival traditions. It was unnecessary for him to remind them that these traditions had grown out of a common stem: that this separation of the churches, now happily reunited, had lasted scarcely 100 years; that recently, as ministers of the gospel, they had tended more and more to forget their differences and to unite in one common purpose. And now, richer in experience and strengthened in faith, they were offered, by the grace of God, a golden opportunity to build up a united church worthy of its name—the Church of Scotland.

"Much, then, has been done," he continued, "and we have good reason to be thankful. And yet in truth our work has scarcely begun. For this Union of ours will be but a sham, a hollow mockery, unless in each and every parish in the land, congregations which hitherto have eyed each other with distrust and suspicion, seek a new, a deeper understanding—forgiving insult and injury, living down hard memories, overthrowing old idols, eradicating those sentiments which lead to dissension and disintegration, and so learning, slowly but surely, to work together in harmony for the

advancement of the Kingdom. And although no two congregations will ever be forced to unite, yet, when opportunity offers, and union seems advisable, the congregations concerned will be asked to unite under one minister who has proved himself acceptable to both. It is those congregations above all others which will have to make sacrifices, to exercise real self-denial, to subordinate their parochial interests to the welfare of the church as a whole, and in general to show true love and forbearance. But theirs will be the greater reward."

Some years later, in Thorness, a town of some 2500 inhabitants on the north-east seaboard of Scotland, the difficulties of putting this ideal into practice became abundantly clear. In this instance, the difficulties had, in the main, two causes—the peculiar temper of the people, and a most remarkable combination of circumstances.

The temper of the people is best indicated by a description of their environment, for they seem to have taken on many of its qualities. To a stranger, Thorness appears as a cold, hard, business-like place, swept clean by the winds off the sea. Most of the houses are built of solid, grey sandstone, and the newer part of the town is well-planned, the streets being broad, well-paved, and set at right angles to one another. The old town, down by the river, where the fisher-folk dwell, is more ramshackle and much less prosperous-looking. These fisher-folk comprise less than one-quarter of the total population. Most of the men are fine, hardy fellows and expert seamen, and the women alert and active, but of late, owing to the failure of the fishing industry, they have fallen on evil days, and poverty and enforced idleness has brought about some degeneration in character: They are never great 'church-goers', although, like fisher-folk elsewhere, they are prone to short-lived bursts of intense religious enthusiasm.

The new town has grown up as a market centre for the rich agricultural district which surrounds it, and its population therefore consists mainly of artisans and middlemen, with a sprinkling of bankers, lawyers, doctors, school-teachers and retired farmers. It is these middle-class people who control church affairs. They are hospitable enough in their way, but just rather than generous, upright rather than amiable; direct in speech as in business; self-reliant, undemonstrative, very 'dour' if crossed. In religion they are moralists rather than mystics, showing little appreciation of, or reverence for, institutionalism or ritual as such. And, as one would expect, they are no respecters of persons. Perhaps they inherit their characteristics from their Norse ancestors, who founded Thorness 1000 years ago, and intermarried with the native Celtic stock. It is at least certain that the early descendants of this union, when oppressed by a worldly-minded bishop, remarked tersely, "Short rede is good rede; slay we the Bishop", and proceeded to roast him in the greater barn which he had built to hold his tithes.

There are seven churches altogether in Thorness, not counting the Salvation Army, the Faith Mission, and a few disconsolate Plymouth Brethren left high and dry by the last tide of revivalism. When the Act of Union was passed, the churches whom it did not concern—the Episcopalian, the Congregational, the 'Wee Frees', and the Original Seceders—waited with a certain cynical interest to see how it would affect the conduct of the other three. They changed their names, of course, the old Established Church becoming St Ninian's Church of Scotland; the first United Free becoming St John's Church of Scotland; and the second U.F. becoming St David's Church of

Scotland. But for a couple of years nothing exciting happened. Each church retained its own minister, St Ninian's remained, as always, somewhat aloof, and one could feel, beneath an outward show of friendliness, the same old undercurrent of jealousy between St John's and St David's. Of the two, the former had always been richer and more vigorous, and there was never any question of its being united with another church. But when the minister of St Ninian's—he was an old man—became weak and ill, and his congregation began to dwindle, as his preaching lost its power, the Presbytery decided that, on his retirement, every effort should be made to unite St Ninian's and St David's. The Convener of the Committee on Union pointed out that St Ninian's would be invigorated spiritually; St David's (and the Church of Scotland as a whole) would benefit economically. St Ninian's possessed one of the finest church buildings in the north of Scotland, well capable of housing both congregations; the church of St David's was small, and falling into disrepair, and those who worshipped there were too poor to restore it. It was true, of course, that neither church had ever shown an active neighbourly interest in the other. Yet at least they did not openly dislike one another, and the minister of St David's had expressed his willingness to be transferred, if necessary, to another charge, so that all the old associations might be broken off, and his congregation given every chance to develop a real unity in loyalty to a new minister. The cynics whispered that this was just another attempt on the part of the former U.F. ministers—they had a large majority in the united Presbytery—to bolster up one of their weaker congregations by allowing it to fasten on the wealth of the older church, and they predicted that the new minister would be an 'ex-U.F.' man, rescued from some hungry parish in the Highlands. But in spite of its detractors, the scheme would probably have gone through smoothly enough—perhaps a little lackadaisically—if it had not been for the extraordinary series of events already mentioned.

What happened was this. At the request of the minister of St Ninian's, who was due to retire in six months' time, a young man, MacKeen by name, was sent north from Edinburgh to help him in his preaching and in his parochial work generally. The appointment was understood to be a temporary one, so no very exact inquiry was made into the new assistant's credentials. His first sermon made a tremendous impression on a meagre handful of loyal worshippers; his second, preached in the evening of the same day, was, as the local papers put it, "much appreciated by a well-filled church". On the following Sunday those who came late had to stand, for all seats were taken. His width of appeal was amazing. Not only did staid middle-class Thornessians wax eloquent in his praise, but the fisher-folk from 'down the town' came over and over again to hear him, and people from the surrounding countryside who had no connexion whatsoever with the church could be seen hurrying townwards every Sunday morning—in cars, on bicycles, in 'buggies', on horseback and on foot. Not only could Mr MacKeen preach popular sermons, he could also sing. Every Sunday evening his congregation was "favoured with a sacred solo from the pulpit"; and seldom did a week pass but he penetrated to the remote fastnesses in the interior of the country, there to win encore after encore on some rustic concert platform. Some of the shrewd old elders of St Ninian's were none too sure of their man, distrusting especially his love of publicity, but they held their peace. The end of it was that the revitalized congregation petitioned the Presbytery that St Ninian's should be allowed to retain its independence (there being literally no room

for another congregation) and that Mr MacKeen should be appointed assistant *and successor* to the old minister.

Whereupon the Presbytery had to take the matter seriously in hand. They found that owing to an oversight on the part of the authorities in Edinburgh, and a lamentable slackness in their own procedure, they had let loose in St Ninian's a man who, by the laws of his church, was not even qualified to be an assistant, much less to be ordained as a minister in full charge of a parish. They found other things also—a story of failure as a student and brilliance as a preacher; of numerous departures from one part of the country to another; of a short-lived blaze of popularity in each district and a willingness to be employed by a church of almost any denomination. Partly out of a desire to spare MacKeen, and partly in order themselves to avoid dangerous publicity, they rested their case on the technical point alone, and 'with much regret' asked MacKeen to resign.

Practically the whole town of Thorness rose up in arms against this verdict. "We have no time", they said, "for ministers of the gospel who hinder the preaching of the gospel." Beyond this they could not, or would not, see, except to speculate as to the real motives behind the decision, for they rightly suspected that so much good work would not be cut short, and so much ill-feeling aroused, for the sake of a mere technicality. The more outspoken of them had much to say of meanness, jealousy and self-interest, and the noise of the dispute was heard far beyond the borders of the country. Newspaper men hurried to Thorness; Mr MacKeen interviewed them all: and the Sunday papers revelled in this "Terrific Religious Crisis in the Far-flung North."

Its climax was indeed dramatic. On the Sunday before that on which he was due to resign, Mr MacKeen announced his intention of preaching a farewell sermon on the following Sunday. The Presbytery asked him not to do so, and they requested a minister from a neighbouring parish—one McTavish by name, an oldish man, who'd been a bit of an athlete in his day—to appear in the pulpit of St Ninian's. On that famous morning a sight was seen which will never be forgotten by those who witnessed it—two ministers racing for the same pulpit, and the older man getting there first by a good five yards! Mr MacKeen took his place in the congregation. The church was, of course, full to overflowing, and a curious crowd was straining round the doors. Mr McTavish rose to announce the first psalm. Before he had opened his mouth, Mr MacKeen rose also, and signed to his congregation to do the same. Then, with a look of Christian resignation on his face, he led the way out. Some of his rowdier followers as they went began to curse the Presbytery, the church and all that it contained, but Mr MacKeen rebuked them, and they filed out in silence. When the last of them had gone, Mr McTavish's voice echoed emptily over the heads of a dozen worshippers—the shrewd old elders, and three or four old ladies who were too rheumatically to move.

After that the Presbytery could afford to let things take their course. They saw MacKeen hire a hall, and set himself up as an independent preacher. They saw him retain his popularity for another six months. Then little by little he began to lose his support. The zest of battle was gone, and the feeling of being martyrs in a common cause soon loses its binding force when one's self-denial gets no publicity. So first of all the fisher-folk lapsed into their usual Sabbath lethargy, and the country people lost their enthusiasm for Sunday travel. The poorer families found it very expensive to have to support a church of their own, and both they and the wealthier citizens

among MacKeen's supporters could scarcely close their ears to the ever-increasing stock of rumours concerning their minister.

The Presbytery waited nearly a year. By that time most of those who had once been staunch members of St Ninian's had deserted MacKeen, and were, so to speak, sitting on the spiritual fence. The Presbytery succeeded in finding a young man (trained in the now "*United Theological College*") who was likely to make a sound and lasting appeal to any decent congregation. He was asked to stand as candidate for the united church of St David's and St Ninian's, the former minister of St David's having been, at his own request, transferred to another parish. St David's accepted him gladly: what was left of St Ninian's did so also, for any state was better than their present one. The formalities of union were completed, and the congregation of St David's entered boldly into the almost empty church of St Ninian's. The new minister, on being introduced to his united congregation, referred tactfully but firmly to the recent regrettable dispute between laity and regular clergy. No blame was apportioned on either side, but he hoped that, out of apparent evil, good would come, and that out of disruption a deeper and more lasting unity would be born. Perhaps in the course of time he would win back to St Ninian's not only those right-minded people who had been so strangely misled, but also some of those who had found their way there temporarily, and were now churchless once again.

Already, three months after his appointment, he has gone a long way towards realizing his ideal. The 'right-minded people' who had been sitting on the fence slipped easily into the fold, where they had the pleasure of reclaiming the best seats, and of reasserting their place in church politics. Then MacKeen, for reasons into which we need not enquire, had to leave Thorness; but even before he went many of his supporters had gone over to St Ninian's, preferring the solidarity of organized religion to the instability of an isolated church. A keen yachtsman himself, the new minister handles his fisher-folk as skilfully as he handles a boat, and they are now developing an interest in religion so well-grounded that it seems likely to be permanent. And in fact good has come out of evil, for the recent turmoil seems to have purged the town of strong emotion, the congregations of St David's and St Ninian's have so far fused with no wild outbursts of feeling, and little or none of the bickering and petty jealousies which so often make true co-operation impossible.

This is the longest and most complex example so far used. In writing it, I endeavoured to give a narrative of 'plain fact', any 'explanation' of the sequence of events being thrown in as a suggestion rather than as a true interpretation. The somewhat high-sounding 'Moderator's speech' is added for two reasons—to give historical perspective and (more important) to place in juxtaposition the real and the ideal, and to see whether in explaining what had occurred, and in predicting what was likely to occur, any of the selected groups of subjects would stress the high motives set forth in the speech.

(ii) *Method of presentation*

This example has so far been presented to 100 subjects. These fall into five groups:

- English: *Academic*: 25 University students, research workers and lecturers.
 Non-academic: 15 'W.E.A.' pupils, mostly engaged in practical, wage-earning activities.
- Scottish: *Academic*: 20 University students (75 % studying psychology).
 Non-academic: Clerical: 20 ministers of the Church of Scotland in various parts of the country.
 Lay: 20 middle-class Scots, nearly all 'church-goers'.

The average age and age distribution in the English groups are closely similar; in the Scottish groups, the average age of the clerical group is rather higher than that of the lay, and much higher than that of the academic.

With the exception of about 50 % of the English academic, and 75 % of the Scottish academic, all replies were obtained by correspondence. The English exceptions dealt with the material in individual interviews, and the Scottish as a group.

In all three modes of presentation, each subject was asked to read through the material presented, and to refer to it as much as he pleased in answering the following questions:

- (1) What factors do you consider were of most importance in bringing about the union between St Ninian's and St David's?
- (2) Do you think that this union will be permanent?
- (3) Do you think that this union will be harmonious?
- (4) Is it likely that the fisher-folk will remain loyal to the united church?
- (5) How would you explain the rapid formation of an enthusiastic group of followers round MacKeen, and the subsequent disintegration of this group?

It was carefully pointed out that if any subject found himself unwilling to express a definite opinion he should say so.

III. ANALYSIS OF RESULTS OF CASE I

(1) *Numerical analysis*

I propose first to give, in Tables I–III, a simple numerical analysis of the answers to questions (2), (3) and (4), and then to attempt to classify, and to determine the relative emphasis on, the leading conditions on which the various types of answer to all five questions have been based.

Table I. *Will the union be permanent?*

Group	Yes	No	Conditional	Total
English: Academic	17	2	6	25
Non-academic	9	2	4	15
Scottish: Academic	11	9	—	20
Non-academic: Clerical	15	2	3	20
Lay	19	—	1	20
Total	71	15	14	100

Table II. *Will the union be harmonious?*

Group	Yes	No	Conditional	Unwilling to express opinion
English: Academic	10	5	10	—
Non-academic	3	3	—	9
Scottish: Academic	4	12	1	3
Non-academic: Clerical	12	3	5	—
Lay	16	—	4	—
Total	45	23	20	12

Table III. *Will the fisher-folk remain loyal to the united church?*

Group	Yes	No	Conditional	Unwilling to express opinion
English: Academic	—	12	6	7
Non-academic	—	6	9	—
Scottish: Academic	—	17	3	—
Non-academic: Clerical	1	9	7	3
Lay	—	9	11	—
Total	1	53	36	10

Note. Of the 'Noes', the following in each group predicted loyalty as long as the new minister remained, then final lapse: 8, 4, 5, 0, 5.

With regard to these three tables, the following points seem worth noting:

(1) No subject in any group found himself unwilling to hazard an opinion as to the permanence of the union, although 14 % altogether give conditional answers—25 % of the two English groups and 6.6 % of the three Scottish groups.

On harmony, the total percentage of conditional answers goes up to 20 %, and 12 % find themselves unable to express an opinion. This 12 % is made up from two groups only, in the following proportions—60 % of the English non-academic and 15 % of the extremely ‘canny’ Scottish academics.

In Table III, the percentage of conditional answers goes up to 36 %. The percentage of those unwilling to express an opinion remains about the same, but is made up from two different groups—English academic and Scottish clerical.

These figures seem to bear out very forcibly Bartlett’s contention that, “when an ordinary person is confronted with a concrete description of a social situation . . . he can usually give a considered opinion as to its probable upshot” (*loc. cit.*). Indeed, in view of the paucity of data on which to frame an answer to questions (3) and (4), it is surprising that more subjects did not refuse to answer. Conditional answers are, of course, almost as valuable as a direct affirmative or negative, for they generally indicate what leading conditions the subject has in mind, although particular determining factors are considered absent.

(2) While 71 % think permanence likely, only 45 % predict harmony. In other words, 26 out of 71, or 37 % of those who predict permanence, evidently consider that a certain amount of disharmony will not destroy the union. Of this 37 %, the 50 % who, by giving conditional answers, entertain the *possibility* of disharmony, thereby admit that disharmony and a permanent union are compatible.

(3) Each of the five groups is so small that further use of the material may significantly alter the results so far obtained. With this reservation, the following comparisons are made:

Table I. The distribution of opinion is most similar in the English academic and Scottish clerical groups. The greatest disparity occurs, not between any one of the English groups and any one of the Scottish groups, or between the two English groups taken together, and the three Scottish groups taken together, but between two of the Scottish groups—Scottish academic and Scottish lay. In the former, 55 % predict permanence, and 45 % breakdown; in the latter, 95 % predict permanence, and none breakdown. Whether this disparity has any significance or not will be discussed when the grounds on which these different opinions are based come to be analysed.

Table II. On the question of harmony, the Scottish groups are much more decisive than the English, 40 % of the English academic giving conditional answers, and 60 % of the English non-academic refusing to

answer at all. Taking the two English groups together, 32·5 % predict harmony, and 20 % disharmony. In the three Scottish groups, 53·3 % predict harmony, but on the other hand 25 % think disharmony likely. This is entirely due to the optimism of the non-academic groups being offset by the peculiar pessimism of the academics.

The answers to this question, of course, unlike those to the question of permanency, are conditioned by the connotation which each subject gives to the word 'harmonious'. None the less, it is interesting to note that the subjects who have had most experience of situations of this type in real life are most decisive in their answers. One would have expected that acquaintance with the number and complexity of factors which determine social 'harmony' would have engendered greater caution.

Table III. The solitary cleric who thinks it probable that the fisher-folk will remain loyal to the church does so because "St Ninian's is the church of their fathers; the endowed church, the church of the poor in past times, making no great demand on their pockets and attachment to it being comparatively easy". None of the rest of the clerical group considers that the fisher-folk will remain loyal even to the new minister, although this opinion is expressed by 25 % of each of the other two Scottish groups, and by about 30 % of each of the English groups. The Scottish academics (85 %) are most emphatic in predicting a final lapse, while the opinion of the other four groups is rather similar—from 40 to 48 % in each considering permanent loyalty unlikely.

(2) *Qualitative analysis*

The leading conditions, on which the various answers to all five questions were based, are classified in Tables IV–XII. This classification unfortunately robs the answers of all their 'individuality', but on the other hand seems to give a fairly good indication of the relative emphasis placed on each type of factor both by all five groups taken together, and by each group taken separately. The method of scoring employed was simply to count the occurrence of any factor whatsoever in any answer as one (1 unit), to classify it, if possible, under an appropriate type, and then to add up the total number of 'units' of each type. This explains why the number indicating the emphasis placed by a group on reasons of a certain type is sometimes greater than the total number of individuals in that group. In his "Preliminary Report and Suggestions" (*loc. cit.*) Bartlett determines the emphasis on any one type of factor simply in

relation to the total number of individuals in the group in question. When the method of scoring just described is employed, however, the most accurate index seems to be obtained by calculating the emphasis in relation to the *total number of reasons mentioned by the group*.

(i) *Initial union*.

Table IV gives the classification of, and relative emphasis on, the factors selected as being of importance in *bringing about* union in the first instance. Table V shows the relative emphasis, calculated as described above (i.e. in relation to total number of reasons mentioned by each group separately, and by all five groups together).

In several of our experiments on 'social thinking' we had noted the recurrence of a 'conventionalized principle' which might be expressed as follows—"It is not so difficult to preserve an agreement between two

Table IV

Factor	Group					Total (100)
	English		Scottish			
	Non-		Academic (20)	Clerical (20)	Lay (20)	
	Academic (25)	academic (15)				
I. By uniting, both congregations at least suffer no loss of prestige, and in some respects flatter their self-esteem	17	20	10	34	30	111
II. Union benefits both churches:						
(a) Economically	9	7	4	10	5	35
(b) 'Spiritually'	—	—	—	6	—	6
III. Natural 'conservative' tendency to unite against disruptive factors	8	2	11	—	3	24
IV. Enlivening and cathartic effect of recent turmoil	3	—	25	1	9	38
V. Decline of interest in religion generally (union a desperate measure to survive)	4	—	—	—	1	5
Total no. of 'factors'	41	29	50	51	48	219

Table V

Factor	Group					All 5 groups Av. %
	E.A. %	E.N.A. %	S.A. %	S.C. %	S.L. %	
I	42	69	20	67	63	51
II (a)	22	24	8	20	10	16
(b)	—	—	—	12	—	3
III	20	7	22	—	6	11
IV	7	—	50	2	19	17
V	10	—	—	—	21	2

groups as to bring them together in the first instance; and a breakdown of the preliminary negotiations is more likely than a breakdown of the final agreement". An opportunity to comment on the peculiar chain of events leading up to church union in Thorness was therefore given, in the hope that further light would be thrown on this interesting generalization.

The outstanding characteristic of the replies thus obtained is that the majority of them indicate a further and much more fundamental principle. Explicit reference to this principle *as a principle* is made by only two subjects. It is most clearly enunciated by one of the Scottish clerical group: "Groups are seldom, if ever, capable of disinterested conduct, and the co-operation of groups depends on how far the interest of the group is furthered thereby and the pride of members in their own group, which flatters their self-esteem, satisfied." In nearly every answer, however (except those given by the Scottish academic group), the specific influences selected as being of most importance in the events leading up to union are such that their full significance is revealed only when the principle itself is discerned, and they are seen in relation to it. It seems reasonable to infer, therefore, that subjects who make this implicit or unconscious use of the principle have had, in real life, empirical evidence of its working, and are able to recognize factors relevant to it, although they may never have consciously formulated the principle as such. (In Aristotelian language, they have perceived the 'that', but not the 'why'.) If this conclusion is justified, our enquiry into the way that people think about social situations is not only revealing the nature, frequency and distribution of opinions, but is also bringing us very close indeed to the real causes which govern social harmony and discord.

The term 'interest' in the principle as enunciated is not intended to refer only to *economic* interest, but to include all 'gains'. In order to bring out clearly that the 'prestige-gain' was considered much more important than the purely economic benefits of union in this instance, I have given each a separate heading in Tables IV and V; and this is further justified by the fact that only the economic gain is explicitly mentioned in the text. The 'benefit of being spiritually invigorated', if one might call it such, occupies a rather peculiar position and will be discussed later.

The specific circumstances which I consider to have been picked out, consciously or unconsciously, as concrete manifestations of the 'prestige-principle' at work, are perhaps worth mentioning in detail:

(a) St Ninian's have 'asserted their independence'; by uniting, and 'saving the church'; can return without *loss of dignity*; are returning to *their own* church; will be friendly to St David's, and not stress their claim to priority, because St David's *came to them*; stragglers will return once they see the church is a 'going concern', association with which gives social standing.

(b) St David's have declined gradually, not catastrophically; are 'sacrificing their independence to save St Ninian's'; have a major share in the appointment of the new minister, after the *voluntary* retirement of their own.

(c) The new minister is 'neutral'.

No one subject, of course, mentioned all these factors. Detailed analysis reveals, however, that they are stressed most evenly in the Scottish clerical and lay groups, and in the English non-academic group. The two academic groups pick out as most important the fact that the new minister is 'neutral'—a fact which the clerics themselves completely ignore.

Some notes on the distribution of emphasis can now be conveniently made (Table V):

(a) The emphasis on the 'prestige-principle' is strikingly similar in the three non-academic groups. It is considered rather less important by the English academic, but is still the factor which they emphasize most. Only in the Scottish academic group does it occupy a subordinate position. The probable reason for this will be seen when we discuss Factor IV). Taking all five groups together, we see at once that the 'prestige-principle' is by far the most important.

(b) Considerable weight is placed on the economic gain by the two English groups and by the clerics; the Scottish academic and lay agree in roughly halving this emphasis.

(c) That the clerics alone should stress the benefits of spiritual invigoration is interesting. It should be pointed out, however, that their terminology—"the people's joy in seeing their church full, and in being carried forward on the tide of a deep enthusiasm", etc.—admits of a cynical interpretation in terms of the prestige-principle. Perhaps that would be unfair. But there is one fact which is not brought out in Tables IV and V, and which is important in this connexion. 20 % of the clerical group, and about the same percentage of the total number of subjects, go out of their way expressly to state that the causes which facilitated union in the first instance, and on the basis of which permanent co-operation can be predicted, are very far removed in tone and

motive from those outlined in the Moderator's speech. Perhaps this interpretation is not meant to be extended beyond the facts of this particular case, and may be due to the manner of its presentation. None the less, the point is made very forcibly in the form of a generalization by one of the clerical group—"If you surcharge a congregation with an emotion like love or hate, or any variety of either, you will get it to unite with anyone or anything from 'Mammon' to the 'father of lies'. I have not heard of two congregations, in the Presbyterian period, who united on the basis of loyalty to the Lord Jesus, brotherly love, or social happiness."

(d) A natural conservative tendency to unite against disruptive factors is most clearly discussed by the two academic groups. This tendency is traced to the preponderating influence of the 'solid elements in the community', who will desire to resume the even tenor of their way after the recent emotional outburst, and will be glad to wrap themselves up once again in the 'comfortable conformity of organized religion'.

This opinion may be due either to the description of the "temper of the people" in the text, or to the fact that the majority of those who express it are academic psychologists, or to both these causes.

(e) What I have called the 'enlivening and cathartic effect' of the recent turmoil is stressed to a remarkable degree by the Scottish academic group. The dispute over MacKeen is considered "to have shaken the people out of their lethargy, and to have let them work off 'emotional steam'"; to have surcharged them with emotion, and so fitted them for union (cf. (c) above); to have quietened the die-hards, who in ordinary circumstances would have held things up; to have humbled many by showing them how they have been 'duped'; to have engendered all-important *individual* friendships, and to have accelerated the swing back towards orthodoxy once the inevitable reaction had set in.

It will be noted, in Table VIII below, that many of those who considered the union to have been thus 'emotionally engendered' also consider, for the same reason, that it will not be permanent.

(f) It would be interesting to discover the prevalence in other groups of the generalization implicit in Factor V, which is considered to be of some importance only by the English academics. The generalization might be expressed thus: "Any two groups who fulfil much the same function in a community and who, while vigorous, would have resisted all attempts at union, will co-operate immediately that they are so reduced in numbers or vitality as to be faced with destruction."

(ii) *Permanent union.*

Table VI gives the classification of, and relative emphasis on the grounds on which a *permanent* union is predicted; and Table VII shows the same emphasis in percentages.

Table VI

Grounds	Group					Total 71/100
	E.A. 17/25	E.N.A. 9/15	S.A. 11/20	S.C. 15/20	S.L. 19/20	
I. Relationship to superior groups	7	—	2	—	4	13
II. Factors relieving early 'stresses and strains'	20	15	12	18	19	84
III. Economic benefit	5	7	4	14	6	36
IV. No conflict <i>in principle</i>	1	—	4	10	18	33
V. Stabilizing factors which will soon begin to work	11	9	9	15	11	55
VI. Disinterested desire for welfare of the Church	—	—	—	7	—	7
Total no. of reasons	44	31	31	64	58	228

Table VII

Grounds	Group					All 5 groups Av. %
	E.A. %	E.N.A. %	S.A. %	S.C. %	S.L. %	
I	16	—	7	—	7	6
II	46	48	38	28	33	37
III	11	23	13	22	10	16
IV	2	—	13	16	31	15
V	25	29	29	24	19	24
VI	—	—	—	11	—	3

As might be expected, there is a considerable overlap in Tables IV and VI, for many of the causes which facilitated union in the first instances are also thought to be causes which will make for permanence. The economic benefit is obviously considered to have binding as well as attractive power; and in those factors which are classed together as "relieving early 'stresses and strains'", and which receive greatest emphasis in every group, it is possible to discover the 'prestige-principle' once again, in slightly different form. The absence of factors damaging to the self-esteem of either group, and the presence of some likely to flatter it, is noted; but most subjects derive their confidence, in predicting permanency, not from knowledge of this fact merely, but from the fact that the new minister is *tactful*, knowing the people's 'touchy points', and therefore likely to keep in abeyance factors damaging to their self-

esteem. The Scottish clerical and lay groups make most of this point, but it also receives considerable attention from the English academics.

What was called, in Table IV, the 'cathartic effect' of the recent dispute also comes under heading II, in Table VI, for a few subjects in every group credit the people of Thorness with some genuine repentance (at least on the part of the prodigals) and a desire to make amends.

This relief of early stresses and strains is evidently considered important as it gives the various stabilizing factors (classed under V) a chance to begin their work. These factors are ranked second in importance by all groups except one, and will be discussed in more detail later.

'Relationship to superior groups' takes two forms:

(a) The position of the united church as a sub-group within a larger organization with central control is stressed by the English academic and by the Scottish lay groups, but in different ways. The English academic group are evidently utilizing a more or less 'theoretical' generalization; the Scottish group point, somewhat bitterly, to the fact that the Church authorities have the power, through the imposition of 'economic sanctions', to force congregations to unite, and to make any attempt at disruption almost impossible; and quote examples to show that, in spite of the promise in the Moderator's speech, this power has already been exercised.

(b) The English academic group, in common with the Scottish academic, consider also that the authority of the Presbytery, the body in immediate control, has been enhanced by its tactful handling of the MacKeen case. Hence the emphasis on factors of type I by the English academic group.

(3) It is worth noting, as an interesting sidelight on the method of scoring adopted, that the emphasis on type II factors is relatively low in the two Scottish groups, not because these factors are mentioned *less often* by these groups, but because other factors are mentioned *more often*.

(4) Economic benefit is ranked third in importance by all five groups together, and (strangely enough) receives most emphasis, relative to other factors, in the clerical group.

(5) The Scottish groups point out that "interdenominational antipathies, as between the former U.F. and Established Churches, though still real, are faint, and not based on any *principle*". It is therefore confidently expected, especially by the lay group, that they will, in this instance, fairly soon disappear.

(6) The 'stabilizing factors' (V) correspond in some degree to the 'conservative tendency' noted in Table IV. Some subjects say baldly that 'time and habit' will bind the two congregations together; others are more specific. "People of the Thorness type being naturally conservative", will, it is thought, stick to the habit of attending the united church, once that habit is acquired, and will prefer the "solidarity of organized religion" to any further attempt at disruption. Pride in the new group will be quickly engendered, and the two groups will boast of "getting on together better than they ever thought possible", and the fine church-building of St Ninian's will become a "centre of approving sentiments". The rise of a new generation who have never known the old antipathies will finally seal the union.

(7) Only the clerics do the people of Thorness the credit of suggesting that they *may* be actuated by a disinterested desire for the welfare of the Church. Over against this we must set the statements noted above, in Table IV, II (b).

(iii) *Impermanent union.*

Table VIII

Ground	Group					Total (15)
	E.A. (2)	E.N.A. (2)	S.A. (9)	S.C. (2)	S.L. (0)	
I. Union in reality externally imposed, and as such cannot resolve fundamental differences in religious temperament and outlook	2	—	4	2	—	8 (27 %)
II. Embittering effect of coercive action by Church authorities	—	—	—	2	—	2 (7 %)
III. Chastening effect of MacKeen incident will wear off	—	2	14	—	—	16 (53 %)
IV. Disturbing effect of presence of third church (St John's)	—	—	—	2	—	2 (7 %)
V. Little real religious feeling behind union	—	—	2	—	—	2 (7 %)
Total no. of reasons	2	2	20	6	—	30

Table VIII shows the grounds on which breakdown is predicted. Only two reasons are given any weight, and the others are stated in detail merely to show the type of factor which tends to be introduced the more closely subjects are acquainted with the situation, or kind of situation, presented.

The remarkable emphasis placed by the Scottish academic group on III above is the natural corollary of their explanation of the event of union in the first instance. The causes which brought about co-operation

being, in the opinion of this group, essentially emotional and transitory, they predict that the chastening effect of the recent turmoil will wear off before the two congregations have had time to settle down; that old and new jealousies will then appear, fostered by 'sticklers' and 'die-hards'. In this instance it would be necessary to know only the number and influence of these in order to determine just how soon the inevitable split would occur.

The other ground on which breakdown is predicted is a generalization which one of Bartlett's 'descriptions' ("Strange Effects of Drought", *op. cit.*) had already elicited—viz. that religious differences, being deep-seated and the outcome of inescapable temperamental characteristics, cannot be resolved by any externally imposed union. Being "inherent in human nature", these temperamental variations have in the past "disrupted great bodies of men into parties; and when, after a period of years, the parties try to come together, or are united by others, the antipathies or differences in temperament make themselves evident and resist easy union". I quote the more emphatic of the two clerics who predicted breakdown, and who both, as one might expect, base their generalizations on a study of Scottish ecclesiastical history.

The Scottish academic group, who consider the effect of the recent dispute too superficial to uproot old jealousies, or to prevent the appearance of new ones, do not trace these to any essentially psychological cause, but merely state that they will be resuscitated and fostered by the more conservative and cantankerous elements in the community. It would be interesting to know how far this group, and popular opinion generally, would be prepared to accept the generalization quoted above as an explanation of the undoubted fact that such jealousies and antipathies occur with great frequency in church politics.

The conditional answers do not require to be tabulated. In order to give a definite answer about 60 % of the fourteen subjects would require to know whether the new minister "is really fit to handle the united church"; the other 40 % would require to know the number and probable attitude of influential people in each congregation taken separately. On the whole, however, all these subjects seem to tend towards 'permanency'.

(iv) *Harmonious union.*

Table IX shows an analysis of the causes considered likely to make for *harmonious* relationships; it is practically self-explanatory, in view of the comments which have already been made on Table VI. The general opinion, most marked in the Scottish clerical and lay groups, is

Table IX

Cause	Group					All 5 groups (45)
	E.A. (10)	E.N.A. (3)	S.A. (4)	S.C. (12)	S.L. (16)	
I. Early stresses and strains relieved by:						
(a) Character and tact of new minister	5	3	1	10	10	29 (39%)
(b) Emotional purge of MacKeen incident and some genuine repentance	4	2	2	2	4	14 (19%)
II. Once difficult first period past, risk of disharmony diminishes owing to:						
(a) Stabilizing factors	3	—	2	4	6	15 (20%)
(b) Weakening of interdenominational antipathies (no conflict in principle)	1	—	1	6	5	13 (18%)
(c) Absence of class and economic distinctions	2	1	—	—	—	3 (4%)
Total no. of reasons	15	6	6	22	25	74

that the character and tact of the new minister will be the main determining factor in relieving early stresses and strains. Once the difficult first period is past, risk of disharmony will decrease owing to the various stabilizing factors which also make for permanency, and which have already been discussed.

The only marked difference between the English and Scottish groups is that the former place some faith in the absence of marked class and economic distinctions, while the latter utilize again their special knowledge of church history to emphasize the absence of any conflict *in principle*.

(v) *Union with disharmony.*

Table X gives an analysis of the causes considered likely to produce disharmony. This table also requires little explanation, as there is a

Table X

Cause	Group					Total (23)
	E.A. (5)	E.N.A. (3)	S.A. (12)	S.C. (3)	S.L. —	
I. Differences in outlook and temperament	2	2	1	2	—	7 (20%)
II. Chastening effect of MacKeen incident wears off, and risk of disharmony <i>increases</i>	2	2	18	1	—	23 (65%)
III. Sea-board communities proverbially distrustful and quarrelsome	—	—	2	—	—	2 (6%)
IV. No real religious feeling behind union	2	—	1	—	—	3 (9%)
Total no. of reasons	6	4	22	3	—	35

marked similarity between it and Table VIII. The outstanding point of difference is the increased emphasis on the probable recurrence of conflict once the superficial causes which are supposed to have removed it begin to lose their effect, and every slight disagreement is exaggerated "as such disagreements always are in small and otherwise dull communities".

Of those who refuse to be definite, the Scottish groups specify only one leading condition on which they consider harmony to depend—the attitude and number of influential people, and of 'die-hards', in each congregation; the English academic group consider social harmony is too dependent on particular circumstances even to admit of a prediction; while the English non-academic group give no reason for not expressing an opinion.

(vi) *The fisher-folk.*

The question of the loyalty of the fisher-folk was introduced in order to see how definite an answer can be elicited from a minimum of data. The relevant data in the text consist of three statements—that the fisher-folk are poor, and form almost a separate community within the town; that they are not "great church-goers, but are prone to short-lived bursts of intense religious enthusiasm"; that they were attracted first by MacKeen, and again by the first minister of the united church, who has himself an interest in the sea. As will be seen, the phrase "short-lived bursts of intense religious enthusiasm" seems to have been most heavily 'weighted' by every group of subjects; and to have encouraged considerable elaboration in setting and argument.

Reference to Table III (p. 216) shows that of the fifty-three subjects who think a relapse on the part of the fisher-folk into their usual 'Sabbath lethargy' almost inevitable, twenty-two think that the fisher-folk are likely to remain loyal to the church as long as the new minister is in charge. Every one of the twenty-two gives as his reason the new minister's 'interest in the sea', and evident sympathy with and understanding of the fisher-folk's way of life. The Scottish lay group suggest rather more subtly that their common interest will form an *exclusive* bond between minister and fisher-folk, and so encourage the latter to overcome their self-consciousness in a predominantly middle-class church. None of the clerics considered loyalty to their hypothetical colleague at all likely.

Reasons for predicting a gradual lapse on the part of the fisher-folk into non-church-going habits are given in Table XI. This table brings out clearly the emphasis on the phrase "short-lived bursts of religious

Table XI

Reasons	Group					Total (53)
	E.A. (12)	E.N.A. (6)	S.A. (17)	S.C. (9)	S.L. (9)	
I. Fisher-folk's instability in religion	10	4	12	9	7	42 (59%)
II. Individualistic and conservative outlook	3	2	7	4	2	18 (25%)
III. Poverty (compared with 'bourgeoisie' of Thorness)	1	2	—	4	4	11 (16%)
Total no. of reasons	14	8	19	17	13	71

enthusiasm". Most of the answers go further, and endeavour to explain this instability, and the conservatism of the fisher-folk, as being for the most part due to their way of life, which "produces a temperament and a cast of mind wholly unsuited to consort with a typically urban population". The difference is more than one of economic organization, but the first step towards securing co-operation between two such groups would have to be a change in economic organization—in this instance, the bringing of the fishermen into some kind of 'Fishermen's Union', so as to encourage the abandonment of their highly individualistic and conservative outlook.

Every one of the thirty-six subjects who gives a conditional answer makes the issue turn on the ability and methods of the new minister—will he be able to control class-jealousies and save the pride of both fisher-folk and 'bourgeoisie', and will he be willing to adopt 'MacKeen methods' in preaching? One or two incalculable factors are also suggested—the presence or absence of revivalists, the poverty or prosperity of the fishermen, the possibility of a change in their economic organization, the degree of inter-marrying between fisher-folk and townspeople, and so on. Three of the Scottish lay group suggest ingeniously that, in view of the professional's scorn of an amateur, the minister's interest in the sea may prove a doubtful advantage!

(vii) *The influence of a leader.*

Table XII gives the reasons brought forward to explain the rapid formation of a group round MacKeen.

In the English academic groups 'MacKeen qualities' receive 50 % of the emphasis; in the non-academic groups approximately the same stress is laid on the fact of opposition. Otherwise the relative emphasis in each group is so similar as to indicate either that reasons of the type used are strongly suggested by the narrative presented, or that there is

Table XII

Reasons	Group					Total (100)
	E.A. (25)	E.N.A. (15)	S.A. (20)	S.C. (20)	S.L. (20)	
I. Mere novelty	6	4	6	3	4	23 (9%)
II. MacKeen's qualities	19	13	21	26	18	97 (39%)
III. Action of Presbytery in opposing MacKeen	11	20	19	37	26	113 (46%)
IV. Cumulative excitement	2	2	6	1	2	13 (5%)
Total no. of reasons	38	39	52	67	50	246

widespread agreement about the factors which are important in situations of this kind. Every subject certainly seemed more eager to tackle this question than any of the preceding ones.

MacKeen's qualities are generally thought to have been enhanced by contrast with the old minister, and can be listed as 'youth, personality, leadership, emotional appeal, oratory and unconventionality'. A few of the Scottish subjects think that he must have had some 'genuine message' and several of the clerics point out, perhaps a little enviously, that MacKeen's lack of academic qualifications was more than compensated for in the eyes of the people by the fact that he 'gave results'.

The effect of the Presbytery's action in opposing MacKeen is worked out very elaborately. All groups agree that the opposition, being apparently unjustified, inflamed the pride of MacKeen's followers, and encouraged them to assert their independence. MacKeen himself, on being opposed, immediately became a *rebel* and a *martyr*. Now 'a rebel *within a tradition* is always supported'—a very interesting generalization common to all groups, and evidently intended to apply to rebels 'within' and against *all* kinds of tradition—and as a martyr, 'a good man not receiving justice', MacKeen would attract many to his 'cause'. As one cleric remarks, quoting John Galt—"God's corbies are ill craws to shute at", and it was just when MacKeen, a minister of the Gospel, came to be regarded as being unjustly persecuted, that his supporters became thoroughly consolidated and ready to make sacrifices for him. An essentially psychological but somewhat cynical reason for this 'love of a cause' is suggested by several subjects both English and Scottish: such causes make an eloquent appeal to a man's egotism, and flatter his belief in his own importance.

Table XIII gives the reasons for the disintegration of the group round MacKeen.

Once again the emphasis throughout all groups is so similar as not to merit further analysis.

Table XIII

Reasons	Group					Total (100)
	E.A. (25)	E.N.A. (15)	S.A. (20)	S.C. (20)	S.L. (20)	
I. Mere novelty soon wears off	4	3	3	—	1	11
II. MacKeen's shortcomings	13	12	14	18	15	72
III. Removal of opposition	11	8	12	10	9	50
IV. Financial responsibility	5	6	6	13	9	39
V. Conservatism	10	4	3	7	5	29
VI. Loss of prestige and social status	4	—	2	—	3	9
VII. MacKeen and his church cannot give full 'pastoral services'	—	—	—	5	1	6
Total no. of reasons	47	33	40	53	43	216

The first three reasons are just the antithesis of those postulated to explain the formation of the group. MacKeen is now seen as a personality whose appeal is strictly limited, and unable 'long to oppose the weight of tradition'; while the scandal concerning him in such a small and respectable community can have but one result. The removal of opposition is taken as demonstrating the reverse of the generalization quoted above—"a rebel, though effective 'within' a tradition, becomes ineffective once he sets himself up as wholly independent". The 'cause' loses its glamour and its publicity, and an ignored martyr is merely ridiculous.

(To be continued)

CHARACTERISTIC MOTIVATION IN THE ACTIVITIES OF SCHOOL-GIRLS

BY M. D. VERNON

From the Cambridge Psychological Laboratory

PART II

- III. *Distribution of the drives in the three groups* (pp. 232-237).
- IV. *Interests and occupational wishes* (pp. 238-245).
- V. *Summary* (p. 245).

III. DISTRIBUTION OF THE DRIVES IN THE THREE GROUPS

We have described how the various drives and motivational tendencies functioned in a number of individuals, and have indicated how it was possible to detect the existence of these drives in the responses to questioning. It is now desirable to give some idea of the frequency of these drives and the extent to which they affected the different girls in the three groups. We have already pointed out that it was easier to find examples of certain of the drives in certain of the groups than in others, which indicates that the frequency varies between the groups. If now we study the number of individuals in whom the drives appeared fairly clearly, we find a confirmation of this variation. Table I gives the number of times each drive occurred, and the percentage of individuals in each group showing these drives (the totals are greater than 100, because most individuals possessed more than one drive). These figures must not be taken too literally; it is not possible always to decide exactly what drives each individual possesses. But they give a rough idea of the frequency of distribution. We shall discuss in the next section the exact forms of behaviour in which the drives were manifested.

We observe that the frequency of certain drives was approximately the same in all three groups, while for other drives it varied between the groups. Among the former was *social conformity* (the decrease between Group (c) and Groups (a) and (b) was too slight to give any definite evidence of a decline in frequency). Social conformity was much the most frequent of all the drives, functioning in about three-quarters of all cases. Thus we may infer that the great majority of school-girls from 12 to

Table I

Name of drive	No. of occurrences of the drive in Group			No. of occurrences as percentage of no. in the Group		
	(c)	(a)	(b)	(c)	(a)	(b)
Social conformity	36	22	21	71	76	75
Humanitarianism	10	3	0	20	10	0
Independence	13	6	3	25	21	11
Security	6	5	6	12	17	21
Activity	19	13	10	37	45	36
Superiority	27	11	4	53	38	14
Dominance	4	2	0	8	7	0
Display	8	4	3	16	14	11
Pleasure	10	6	6	20	21	21
Total	133	72	53	262	249	189

18 will seek to make close contact with their social group, to conform to its traditions, imitate its behaviour, or obey its wishes. Any or all of these tendencies indicate the presence of the social conformity drive. In general the social group to which behaviour and ideas were referred was that of school companions. But we have also described cases, chiefly among the younger or emotionally retarded, of conformity to parents; of conformity to school staff; and, among the more mature, of conformity to general social opinion—appearing as a moral obligation towards society. Thus as age increases and emotional maturity is achieved, the circle of society to which the individual conforms becomes wider. But we have described cases in which, for various reasons, the individual is unable to conform freely to the social group to which she should naturally be suited. Emotional backwardness is one such cause; it appeared in those cases who were unable to make contacts with their contemporaries because they were too closely tied to their parents, sometimes because their parents exerted too strong a forcing or restraining influence upon them, keeping them in childish subjection. Intellectual and cultural inferiority also prevented normal contact with and conformity to school companions. Cases such as that of M 5 in Group (c) (this *Journal*, XXIX, 127) will inevitably occur in any social group. They were rather unduly frequent in Group (a) for the following reason: This group consisted partly of girls coming from highly intelligent and cultured homes in the professional classes, and partly of girls from less cultured (and sometimes impoverished) homes in the commercial and business classes. The former were naturally more intelligent and usually showed more energy, initiative and independence than the latter, and consequently tended to dominate the school group and determine its

general social and cultural level. Thus a feeling of inferiority was apt to arise in the girls of lower cultural and social status, and they were unable to conform satisfactorily to the ideas and interests of their more intelligent companions. This inability to conform resulted in shyness and timidity, inability to make social contacts, and, in some cases, the tendency to take refuge in the society of animals. Thus we have included among the cases of social conformity those in whom the drive was thwarted or inhibited, as well as those in whom it functioned naturally and freely.

Activity was also comparatively evenly distributed in the three groups, occurring in between one third and one half of the cases; it was somewhat more frequent in Group (a) than in Groups (b) and (c). The occurrence of this drive was difficult to decide with any certainty, except in some of the outstanding cases described in the last section. In point of fact almost all of these girls were naturally fairly active, except some who were emotionally inhibited. But we attributed the drive only to individuals who seemed capable of fairly purposive and prolonged effort; and not to those who were indeed 'active' in the usual sense of the word, but flitted rapidly from one pursuit to another, with little concentration or perseverance. The more extreme cases of this lack of directedness and perseverance were usually motivated by the drive towards *pleasure* or *variety*. We observe that the relative frequency of this drive was almost identical in the three groups. Thus this type of laziness or superficiality seemed to be an inherent characteristic which varied comparatively little with age or social background.

Coming now to the drives which varied in frequency, we see that humanitarianism, independence, superiority, dominance and display were most frequent in Group (c) and least frequent in Group (b); while security was most frequent in Group (b) and least frequent in Group (c). In all cases Group (a) was intermediate—an interesting point, since the girls in Group (a) were intermediate in age and social status between those in Group (c) and those in Group (b). However, we cannot necessarily infer from these figures that the drives towards humanitarianism, independence, superiority and dominance were almost, or altogether, absent in Group (b). It is possible (1) that these tendencies were too slight to appear in the behaviour of these girls, which was largely governed by conformity to the social norm of the group; and (2) that even when the drives did affect behaviour, most of the girls were too shy and inarticulate to express them in any way in their very brief descriptions of interests, wishes, etc. But against this latter supposition it must

be remembered that the girls' behaviour in school was also observed, and was reported on by their teachers. Thus we may perhaps conclude that certain of the drives do appear more frequently and clearly among girls who are superior in age, intelligence, culture and social status. Which of these factors is the most important cannot always be certainly determined; that almost certainly varies from one drive to another.

To take first the drive of *humanitarianism*: It was pointed out in the previous paper (*Brit. J. educ. Psychol.* (1937), VII, 302 and (1938), VIII, 1) that an interest in people for themselves was comparatively rare among younger girls, and did not appear as a rule until they were nearly adult. The results given in this paper seem to bear out this conclusion. We see that there were few cases except in Group (c) of girls with a genuine interest in or concern for people in themselves—as apart from those who simply liked being with people or working with them. The latter we have distinguished as being motivated mainly by social conformity. But a further difficulty arises here. There were in addition, in all three groups, quite a number of sociable and friendly individuals, with not much independence or initiative, who seemed to show sympathy and protectiveness—manifested, for instance, in a wish to take up work with little children; B 2 in Group (c) (see this *Journal*, XXIX, 126) was a case in point. But it was very hard to distinguish among these cases the genuine humanitarian from the rather childish and unintelligent social conformer, since the latter might wish to do work with children, partly because it seemed to promise her less intellectual strain, and partly because in undertaking such work she could hope for social approval. Thus we have only assigned the humanitarian drive with any certainty to those girls with sufficient independence and initiative to have adopted this interest in people spontaneously, not merely to have followed social tradition. This is not to conclude that there is a necessary positive correlation between independence and humanitarianism (so that because there was more independence in Group (c) there was therefore more humanitarianism); but merely that we can detect the latter with more certainty when the former is also present. At a later age it might be possible to decide which among the social conformers were also genuine humanitarians; many of those who wished to take up nursing or child nursing, or even work with animals, might turn out to be such.

There is little doubt that *independence* was more frequent in Group (c) than in Group (a), and more frequent in Group (a) than in Group (b). Age may again have been partly responsible for this; girls may develop more independence when they have passed adolescence. Favourable and

unfavourable selection were probably more important. In the Group (c) school, the girls with less independence and initiative might have left immediately after taking the School Certificate, and not stayed on in the sixth form. Among the less well-to-do school population, the girls with more drive and initiative, as well as intelligence, obtain special places at a secondary school such as that from which Group (a) was drawn. Those with the least intelligence and initiative remain in a Senior School, such as that which contained Group (b). But more important in Group (a) than the influence of selection was that of home background. Almost all the girls with most independence were the daughters of academic or professional class parents. Thus probably their independence was to some extent inherited; and also every sign of it was approved of and encouraged by parents.¹ It is probable that the same influence was at work in some of the girls in Group (c), although its effects were here less clearly marked. In Group (b), however, independence may quite possibly have been discouraged, both at home and at school.

The influences of selection and of home background were also at work on the distribution of the *superiority* drive. In the Group (c) school the most ambitious girls would tend to stay on at school in order to work for the University and for a profession. Ambition and desire to rise in social status would also lead girls to go to a secondary school, such as that of Group (a), and not remain at a Senior School as in Group (b). Indeed, we saw that two girls in Group (b) who did show the superiority drive were trying to go to the Technical School (they were too old to go to a secondary school). It was also true that the daughters of professional class parents such as those in Group (a) and Group (c) were on the whole the most ambitious, especially in desiring academic success and professional careers. Here indeed parental encouragement and expectations were sometimes so strong as to have a definitely unfavourable influence, since the daughter endeavoured to maintain a standard of performance which was too high for her to achieve without undue strain. A further, less deleterious, influence seems to have been at work with the Group (c) girls. In addition to those coming from professional homes were a large number from well-to-do and upper class homes; these girls were accustomed to a superior social position, and to succeed in obtaining almost anything for which they wished. Thus it was natural that they should expect this success and their superiority to the rest of mankind. This

¹ An exception to this was H 2 in Group (a) (see this *Journal*, xxix, 133); but she had inherited so much independence from her father that it merely became more aggressive when it came in conflict with his commands.

probably accounts for the large number, more than half the group, in whom the superiority drive was observed.

Although the relative amount of *dominance* was greater in Groups (c) and (a) than in Group (b), the actual number of cases was too small to suggest any definite conclusions. It appears that this drive is infrequent even in intelligent girls of superior social status. Probably it is an inherent quality of the individual, not very common in women, and relatively unaffected by social status. But it is also possible that it develops and appears more clearly with increasing age; it occurred in more than a third of the University students (*loc. cit.*), who came from much the same social classes as those of Groups (c) and (a).

The frequency of *display* did not vary greatly in the different groups, though the largest proportion of cases appeared in Group (c). It seems probable that in every social class there are a certain number of individuals who are chiefly interested in occupations and pursuits in following which they may hope to be personally admired. At the same time, this quality may well have been encouraged by the more fashionable and well-to-do parents of the girls in Groups (c) and (a).

Finally, we see that *security* alone of all the drives was relatively most frequent in Group (b) and least frequent in Group (c). It may be supposed that this is in the main a directly environmental effect. Poverty, unemployment and inadequate nutrition were only too familiar among the Group (b) girls. In Group (c) such factors were almost entirely absent except in the two or three cases described in Part I of this paper whose fathers were in relatively insecure financial positions. At the same time it is interesting to observe that the security drive appeared in Group (b) more in a general timidity of manner, a desire to cling to parents and to familiar home surroundings, than in the wish for "a good, safe job" (as in Groups (a) and (c)). The security given by earning a good wage in a job with little fear of unemployment was outside the experience and beyond the expectations of the Group (b) girls; and thus security was constituted by the support of home and parents. On the other hand the Groups (c) and (a) girls were beginning to realize that one can obtain security for oneself, and is not obliged to depend for it on other people.

This concludes the general account of the distribution of the drives; and we shall next consider how they manifested themselves; and observe that the actual interests and occupations to which they lead also varied greatly from one group to another.

IV. INTERESTS AND OCCUPATIONAL WISHES

Since those who were strongly motivated by social conformity endeavoured to imitate the behaviour and adopt the wishes and interests of their social group, it was natural that behaviour, wishes and interests should vary according to the social norm of the group—the traditional and customary ideas, interests, wishes and behaviour—which itself varied considerably with the social setting. In Group (*b*), the group of lowest social status, the norm was the narrowest in range, giving the fewest permitted variations; it consisted chiefly of domestic and practical activities. The main interests were in cooking and needlework in school, and in simple amusements such as reading and knitting out of school. The desired occupation was domestic service. The horizon was bounded by the town nearby and its openings for employment and amusement. Opportunities for physical activity were distinctly inadequate, though singing, physical training, and playing games at school (on a small scale) were usually popular, and sometimes cycling, walking and swimming out of school. Manual activities also provided an outlet for the activity drive. Interest in people was confined mainly to the familiar friends and acquaintances in the school and village; though the wish to do child nursing may have indicated a nascent humanitarianism. The social conformers were fairly contented with this life; but those who possessed a strong security drive tended to shrink away from it, and to wish never to go away from home but live always with their parents in the same familiar places. On the other hand, those who possessed some degree of the superiority drive wished to rise above their social level to a higher one. This could be done by (1) going into a shop or a hairdresser's establishment, or (2) going to the Technical School, training as a shorthand typist, and going into an office—or better still, becoming a school secretary. The first course was suitable for those girls who possessed comparatively little academic intelligence, but some social flair. It also appealed to those in whom the drive towards display existed. The second course was chosen by those with greater academic intelligence. It was probably regarded as the superior of the two since it was likely to lead to a better paid job.

We have already pointed out that Group (*a*) formed in fact a divided social group, one section of girls coming from the professional classes, the other from the business and commercial classes. The latter in the main conformed fairly closely to their social norm, which was not very different from that of the Group (*b*) girls although they were some degrees above

them in prosperity and in social status (the intermediate social classes would go to technical and central schools). This social norm was constituted by wishes and interests which were mainly practical and domestic. It is true that the wider school curriculum gave opportunities which were sometimes adopted of making wider cultural interests—as, for instance, in history, English literature, science, drawing and craftwork; although spontaneous interest in these subjects seemed somewhat inhibited by the feeling of inferiority arising from their more moderate culture and academic ability. Out-of-school interests in reading, knitting, needlework, walking, cycling and swimming were the same as those of Group (b), but playing games, riding, gardening, keeping pets, drawing and music were added, and gave a fuller and more varied scope for activity than was possible for Group (b). Occupational wishes included: secretarial work (4), domestic science (1), nursing (3), work with children (3), work with animals (5).¹ Domestic science and work with children are the counterpart, at a higher social level, of the domestic service (cook and nursemaid) of Group (b). That is to say, the Group (a) girls also desired in the main work of a practical rather than an intellectual nature, which involved plenty of contact with people. The wish to do nursing and work with children possibly indicated the existence of the humanitarian drive; but as was pointed out above, this was not certain. The wish to work with animals was probably a phantasy wish, created by the desire for social security arising in a situation of social insecurity and inferiority. For other girls, security was represented by “a good job”, “something settled”, something in which one can soon earn a living wage. But this desire to be a wage earner was also in some cases a mark of the superiority drive, as, for instance, in the case of H 5 (this *Journal*, xxix, 139). Cases of superiority were, however, comparatively rare in this section of the Group (a) girls. There were a certain number who possessed the drive towards display, which manifested itself principally in a fondness for dancing and acting.

The social norm for the professional class section of the girls in Group (a) was wider, and the permitted variations greater. It was predominantly intellectual in character. School interests were imaginative and cultural, often literary; and three girls had strong scientific interests. Out-of-school interests were numerous and varied, and included artistic and literary pursuits as well as physical and manual activities. Moreover, these interests were usually pursued with more activity and greater energy than by the girls previously described. Occupational wishes

¹ Some of these were alternatives.

included: University careers (7), teaching (4), medicine (1), writing (1), art (1), music (1), the better type of secretarial work (3).¹ It seems probable that the social tradition (supported by parents and teachers) was towards some kind of profession, though secretarial work would be included if as a private secretary or organizer. (The dominance drive also directed people towards this kind of work, as shown by B 2 and T 1.) Nursing seems sometimes to have been suggested but was firmly rejected by at least two girls. The professional careers were in the main those involving work with people. But so many of these girls possessed strong superiority drives; thus they desired in addition (1) a successful University career, (2) an occupation which is "interesting", "not monotonous", something in which "you can use your brains"—in fact, something which gave opportunities of exercising their imaginative and intellectual abilities. But opportunities for social contacts were also considered important, and humanitarianism seemed to appear in one or two of those who, for instance, enjoyed history because it "was about people", and wanted to take up medicine, teaching, etc. Security did not seem to trouble this set of girls.

Although a certain number of the girls in Group (c) came from the professional classes, the social norm was determined in the main by the large number of girls coming from well-to-do homes, whose parents were often fashionable 'society' and 'big business' people. Thus their out-of-school interests were often of the opulent, sporting, games-playing, dancing and theatre-going variety. But they also possessed a wide range of cultural interests. The school curriculum was a varied one, and not all of these girls were doing specialized work; many were carrying out a generalized programme. There were opportunities for studying the usual school subjects, and, in addition, philosophy, economics and politics; for artistic subjects—drawing, modelling, painting, architecture, music, elocution and acting; and for domestic science. Thus cultural and artistic interests were not confined to the more intelligent and original members of the group, but also characterized the social norm to which the rank and file conformed. Particularly noticeable also, throughout the group, was an interest in foreign languages and in travelling abroad. Thus 41% of all the girls were interested in modern languages; 51% had been abroad in their holidays; 37% intended going to live abroad for a period of some months, when they left school; and 16% wanted to take up some occupation in which they could use their languages. This interest seemed characteristic of the social norm in Group (c); it resulted from the greater

¹ Some of these were alternatives.

wealth and wider social opportunities and experiences of the girls in this group.

When we examine the other occupational wishes, we find it hard to determine the nature of the social norm in that respect. It is true that almost all the girls said they wanted a career. But among the less intelligent and independent this wish seemed to be largely a function of the activity drive; that is to say, a definite occupation would give the girl something to do, while if she lived at home without one she might get bored with little or nothing to occupy her. Some thought of taking up nursing or nursery school work (7), though it was not at all certain if they would persevere in these occupations. Others (8) wished for secretarial work, or to go into business. The widespread interests in the arts and crafts were mainly cultural, and did not correspond to natural abilities and inclinations of such magnitude as to make their owners wish to develop them professionally. Table II shows the number of girls interested in these subjects and the number who hoped to take them up as a career,

Table II

Subject	No. interested in subject	No. taking up subject as a career
Drawing, modelling, craftwork, etc.	19	3
Architecture	4	3
Music	21	1
Elocution and acting	16	1
Domestic science	17	2

which is very small by comparison, except in the case of architecture. The choice of architecture may have resulted from a special interest in the subject peculiar to the school, and inculcated by good teaching. But it is probable also that it was adopted because it seemed more likely to provide a secure means of livelihood than would any other artistic profession. The security drive also came clearly into operation in some cases whose parents were not well-to-do, and directed them towards relatively secure occupations such as teaching or the Civil Service, and away from the insecurity of, for instance, acting and writing.

We have already shown that over half the girls in Group (c) were motivated by the superiority drive, and one quarter by the independence drive. The majority of these girls hoped to have a University career and subsequently to enter a profession. Thus in general they were specializing in one or more subjects in school, and they possessed strong academic interests (usually in English, history, modern languages or science), in addition to the more generalized cultural interests we have described

above. The main occupational wishes were: teaching (4), medicine (5), secretarial and business (4), Civil Service (Treasury and Departmental classes) (5), research work (5), social work (5), writing and journalism (5).¹ The Civil Service, writing, research work, were chosen particularly by those who wished to demonstrate their superiority in excelling in their University and subsequent careers, and who also desired to exercise some independent initiative. The same motives appeared in many of those who were taking up medicine and were determined not to become general practitioners when qualified, but to obtain some superior position, out of the general ruck (in surgery, tropical medicine, welfare or clinic work). Journalism seemed a good opening to those with a rather superior and detached general interest in human activities, affording plenty of opportunities for independent initiative, and also plenty of change and variety. Social work and teaching appealed to those with a feeling of moral obligation towards society. This was in part the outcome of a genuine humanitarian feeling for people. But in some cases it also seemed to derive from social conformity—the desire to obtain social approval for a type of work which also gave opportunities for dominance and superiority (see, for instance, M 4, described in this *Journal*, xxix, 136). Secretarial work seemed to be somewhat of a *pis aller* to the individual who possessed strong superiority and independence drives. She regarded it as an occupation more suitable for her less gifted companions, but one which it might be necessary to adopt if other more interesting careers were impossible; thus A 2 said that “if she fails (to get into the Civil Service), will be the inevitable secretary”.

In cases in which there was a drive towards display, there was often great interest in some form of artistic self-expression, such as acting, elocution or dancing, with a wish to take it up professionally. When there was also a strong superiority drive, the individual desired to excel in these professions, or else rejected them altogether, for fear of not excelling. Another occupational wish probably arising from these drives was to go into Parliament. But sometimes display was linked with the drive towards pleasure and variety. In such cases the individual usually wanted to act or dance as a pastime or a part-time occupation, and not as a full-time career. The pleasure seeker was in general characterized less by any particular set of interests and occupational wishes than by a general changeableness, a desire for variety and a lack of perseverance. This might result, as in the case of S 1 in Group (b) in the mainly passive behaviour of staying at home and reading novels; or it might

¹ Some of these were alternatives.

appear as the desire to "have a good time", with plenty of games, amusements and travelling; with plenty to do, but no obligation to work hard either in a career or in housework carried on at home.

In Table III are summarized and outlined the interests and occupational wishes in which the various drives appeared to manifest themselves in the three groups. Group (*a*) is subdivided into (1) those coming from non-professional class homes, (2) those coming from professional class homes. The activity drive has been omitted, since this seemed in general to reinforce the carrying out of activities the nature of which was determined by other drives.

We see that, just as some drives varied greatly in frequency of distribution in the different groups and others very little; so the manifestations of some drives varied more between groups than did those of others. Clearly the manner in which social conformity appeared varied as the social norm varied. The superiority drive showed a like variation, at least in occupational wishes, since its possessor wished to rise superior to the social norm. Display directed the individual towards some occupation which was admired in that particular social setting, varying from that of hairdresser to that of Member of Parliament. Security led to a retreat from the unfamiliar to the most familiar (the home) in Group (*b*); but in Groups (*a*) and (*c*) it also led to positions of economic security or social escape (work with animals). Humanitarianism varied less, though even there in the higher social classes the interests were more academic and the occupations mainly professional. But independence naturally led its possessor to seek for something independent of social tradition, in which originality and initiative could function; and the Arts seemed to form the most obvious vehicle for these. Dominance required roughly the same type of social setting throughout (a leader and subordinates) for its manifestation, though its actual context varied; and indeed opportunities for its exercise were very slight for those of lower social status. Finally, pleasure varied a little in context but was always characterized by changeability and a general lack of effort and perseverance.

Again, it is clear that any one interest or activity may be motivated by very different drives, even in the same social group, and still more in different social groups. Thus teaching may be desired through (1) social conformity, (2) humanitarianism, (3) security, (4) superiority, or (5) dominance, or any combination of these drives. Artistic activities which are normal to the social conformer in one group may be original (independent) or superior for individuals in other groups, or provide opportunities for pleasure or display for others. Clearly then we are not justi-

Table III

Drive from which interests and wishes derive	Interests				Occupational wishes		
	Group (a)		Group (b)		Group (a)		
	(1)	(2)	Group (c)		(1)	(2)	Group (c)
Social conformity	Domestic, manual, physical* activities	Intellectual, artistic, * manual activities	Domestic, artistic, manual,* activities; humanistic interests; foreign languages and travel	Domestic	Secretarial, nursing, work with children	Professional, higher secretarial	"A career"; secretarial, nursing, work with children
Humanitarianism	?History ?Playing with children	?History ?Science	Zoology; history, economics, politics, philosophy	?Nursemaid	?Nursing, work with children	Teaching, medicine	Social work, teaching, medicine
Independence	—	Original work, especially artistic and literary	Original work, especially artistic and literary Dislike of going away to school	—	—	Artistic and literary work; "something interesting,"	Artistic and literary work; journalism; research
Security	Doing things with animals	—	Dislike of going away to school	Staying at home	Secretarial; work with animals	—	Living at home; work with animals; teaching, Civil Service
Superiority	—	Intellectual or academic subjects, especially those at which she excels	Specializing in academic subjects	Technical school, shorthand typist; hairdresser, shop assistant	Being a wage earner	University career, professional success	University career, Civil Service, writing, research
Dominance	—	School duties	Prefect and games duties	—	—	Secretarial, organizing	Social work, teaching
Display	Dancing, acting, games	Acting	Dancing, acting, physical activities	Shop-work, hairdressing	Dancing	—	Dancing, acting beauty culture M.P.
Pleasure	Games, amusements, travel; anything new	—	Games, amusements, travel	Not going out to work	Freedom from parents and domestic work	Something interesting, not monotonous	Having plenty to do, but not a regular job

* 'Physical activities' include: games, sports, physical training.

fied in predicting anything as to the nature of interests, or of the behaviour or occupational wishes to which they lead, unless we know something of (1) the basic drives of the individual who possesses these interests, and (2) the social setting and background from which he comes and in which he is situated.

V. SUMMARY

1. From the answers to questions given by three groups of girls of different ages and social settings, was demonstrated the existence of the following drives towards: social conformity, humanitarianism, activity, independence, security, superiority, dominance, display, pleasure or variety.

2. The frequency of distribution (that is to say, the proportion of individuals in whom the drive appeared clearly) was roughly the same in the three groups for social conformity, activity and pleasure. In the group containing the girls of highest age and social status the drives towards humanitarianism, independence, superiority, dominance and display were the most frequent, and security the least frequent; and the position was reversed in the group containing the girls of lowest age and social status. The intermediate group was intermediate in frequency for these drives. The various factors which produced this effect are discussed.

3. The particular interests and occupational wishes towards which the drives directed individuals differed in the different groups, particularly in the case of social conformity, superiority, display, security and humanitarianism. Moreover, any particular interest or occupational wish might be produced in different individuals, in different social groups, by very different drives. Hence the occurrence of any such interest or wish has no significance unless considered in conjunction with the basic drives of the individual (and also probably his age and emotional maturity), and with his social background and environment.

In conclusion, I wish to make acknowledgement to the Medical Research Council for a personal grant which made it possible for me to carry out this work. My thanks are due to Prof. Bartlett and to others for whose advice and suggestions I am most grateful; and most particularly to the Headmaster, Headmistresses and Staff of the three schools in which these observations were made, for their kindness and helpfulness.

(Manuscript received 11 May 1938)

A PRELIMINARY NOTE ON A VOCATIONAL INVESTIGATION

BY ALICE HEIM

- I. *Introductory* (p. 246).
- II. *Sources and type of data* (pp. 246-247).
- III. *Results* (pp. 247-251).
 - (1) *Employability* (pp. 247-249).
 - (2) *Choice of work* (p. 249).
 - (3) *Personal qualities* (p. 250).
 - (4) *Criteria of satisfaction* (pp. 250-251).
- IV. *Summary* (p. 251).
- V. *Conclusion* (p. 251).

I. INTRODUCTORY

IN this paper, a preliminary survey is made of some of the points which have arisen in the course of an enquiry into the working of the present system by which elementary school-boys enter industry. The Industrial Health Research Board has kindly allowed the investigator a grant, enabling her to carry out the work.

II. SOURCES AND TYPE OF DATA

The main aim of the investigation is to determine the importance of, and the relations between, a number of the factors which influence the early industrial lives of elementary school-boys. The data are drawn from four schools, three Senior and one Central. Material has been derived from the boys' school-masters, relatives, doctors, employers and employment officers, and information obtained as to the boys' scholastic ability, appearance and manner, temperament and character, physique, home environment, choice of occupation, satisfaction and satisfactoriness in work, spare time activities and employability.

The material here discussed is derived from the first of the three groups of school-leavers who are the subjects of the investigation. This group consists of ninety-eight boys, most of whom left school in March 1936. (Some of the Central School boys left later in the same year.) The

investigator gave a group test to all these during their last term; this was a comprehensive test which takes half an hour to give and consists of three parts: Directions, Analogies and Opposites, and Mechanical Aptitude.

The scholastic ability, appearance and manner, and temperament and character of the boys were judged by their school-teachers and recorded on forms constructed on a three-point scale. The first form asked for information such as the boy's speed and originality of thought, imagination and consistency in school work; good and bad school subjects were also mentioned. The second form dealt with such characteristics as normal facial expression, voice, clothes, and the third listed various social, emotional and 'moral' qualities and their opposites.

Full reports on the health and home environment of each boy were obtained from the School Medical Officer and the School Enquiry Officer, respectively. The investigator also obtained some first-hand information on home conditions by personal visits to the boys' parents, when hobbies, ambitions and views on present and past situations were discussed. Data as to choice of work and work obtained were derived also from the Juvenile Employment Office.

An 'employability index' was made out for each boy on the basis of the quality of the situation (or situations) he was in during his first year after leaving school. A fivefold division was adopted, ranging from α , i.e. highly skilled occupations with excellent prospects, to γ , i.e. quite unskilled work with no prospects. When a boy had had more than one situation, an attempt was made to estimate his 'average' employability; in these cases, the upward or downward trend of a boy's career was also taken into account.

III. RESULTS

(1) *Employability*

Fourfold tables were constructed to show the relations between various pairs of the above factors, and the probability of obtaining the observed relations by chance was calculated in each case by the χ^2 method.

The following tables show the connexion between employability and test scores (I), scholastic ability (II), home environment (III), and appearance and manner (IV).

It will be seen that the relationship is in each case statistically significant. It is interesting that the connexion between employability and the test scores should prove to be stronger than that between

Table I. *Relation between employability and test scores*

Employability	Test interquartile groups	
	1 and 2	3 and 4
α , β + and β	26	13
β - and γ	13	24

$$\chi^2 = 7.5. \quad P = 0.006.$$

Table II. *Relation between employability and scholastic ability*

Employability	Scholastic interquartile groups	
	1 and 2	3 and 4
α , β + and β	34	22
β - and γ	13	25

$$\chi^2 = 6.36. \quad P = 0.01.$$

Table III. *Relation between employability and home environment*

Employability	Home environment	
	α and β +	β , β - and γ
α and β +	37	3
β , β - and γ	21	35

$$\chi^2 = 29.35. \quad P = 0.000001.$$

Table IV. *Relation between employability and appearance and manner*

Employability	Appearance and manner	
	V.G. and G.	Av., F.O. and P.
α and β +	27	11
β , β - and γ	20	36

$$\chi^2 = 11.31. \quad P = 0.0008.$$

employability and scholastic ability: the latter distribution would occur as a matter of chance once in one hundred, the former, only six times in one thousand.

However, according to the values of χ^2 found and given above, both home environment and appearance and manner were much more closely related to employability than were the scores for intelligence or for scholastic ability. This would suggest that the importance of the role played by certain social factors in the boys' early industrial lives, although suspected, has perhaps not been fully recognized in the past. On the other hand, it should not, in the opinion of the investigator, necessarily be inferred that a boy's home environment and appearance and manner play in fact a greater part in his industrial career than do his more intellectual abilities and attainments: even if this be true of his

first year after leaving school, it seems likely that this importance will decrease as he grows older. Home influences will probably weaken during the next few years and emphasis may gradually shift from *appearance* to *experience* when vacancies for older boys are to be filled.

(2) *Choice of work*

A comparison of the boys' choices of work with those chosen for them by their parents, suggested that parents are on the whole more ambitious for their sons than the boys are for themselves. Table V shows the distribution of the vocational wishes, within the categories of 'skilled', 'semi-skilled' and 'unskilled' occupations. Office work, shop work, the Services and the railway were listed separately, as they cover a wide range of occupations with considerable variation in the degree of skill required. The figures in brackets at the top of each column give the total number of choices which were ascertained and from which the percentages in the table are taken. Thus, 32·2% of the eighty-seven boys whose first choice is on record expressed a desire for skilled work.

Table V. *Boys' and parents' choice of work*

Type of work	1st choices (87)	2nd choices (70)	Parents' choices (41)
	%	%	%
Skilled	32·2	17·1	31·7
Semi-skilled	23·0	24·3	14·6
Unskilled	9·2	18·6	0·0

It will be seen that while the percentage of parents whose choice is for skilled work is almost identical with that of the boys' first choice, there is considerable difference in regard to 'semi-skilled' and 'unskilled' occupations; not one parent who expressed any views wished his son to enter an unskilled occupation, although 9·2% of the 'first choices' and 18·6% of the 'second choices' are for such work. The hypothesis that those parents who possessed definite and explicit opinions as to their children's careers were also the most ambitious merits consideration. No evidence has been obtained on this point, but it is one which should not be ignored in any discussion of the above results.

A large proportion of parents wished their sons to go on to the railway or to do office work, the former on account of its reputation for security, and the latter largely on account of the social status which is conferred by cleanliness and shorter hours. These occupations were slightly less popular among the boys.

(3) *Personal qualities*

The investigator has called on, and corresponded with, the boys' firms in an endeavour to obtain a report on the suitability, progress and conduct of the boys in their work and also to ascertain employers' general views on juvenile labour. Employers were asked what they considered to be the most important qualities required by employees who were engaged on work such as their firm did. Responses ranged from silence to long harangues, and the subject-matter of the answers given presented an equally wide range. From the answers it appears that employers either do not know or cannot express what it is they look for in their employees. If this is the case, the high juvenile labour turnover found by the investigator in numerous firms (at a time when the number of available positions exceeded the number of available boys) is hardly surprising.

However, an examination of the relationship between temperament and character traits and employability suggests that some kind of sorting-out process, on the basis of personal qualities, nevertheless takes place. When these traits were compared with employability, it was found, for instance, that no boys who were designated as outstandingly *conscientious* have had γ situations. On the other hand, among the α situations, not a single strikingly *lethargic* boy was to be found.

Owing to the paucity of the number so far examined, these findings may be far from reliable. They suggest, however, that further research along the same lines might yield interesting results.

(4) *Criteria of satisfaction*

An analysis of the reasons given for dissatisfaction with work, and for leaving situations, indicated that the physical conditions under which boys work and the type of work which they are required to do play a greater part in determining the degree of a boy's satisfaction, and the length of time he will remain in a situation, than material considerations such as hours and wages. This conclusion is opposed to the views on the subject expressed by most employers and many Juvenile Employment Officers, who usually account for juvenile discontent and unrest in terms of insufficient remuneration.

The above finding, however, is confirmed by two facts which became evident in the course of the investigation: (i) the continuance of a high juvenile labour turnover in various firms which have recently raised their wages while continuing to ignore other relevant human factors; (ii) the

existence, in several of the larger firms, of a higher labour turnover in their highly paid, mass production departments than in their less well-paid workshops, where boys are taught a skilled trade.

Among elementary school children, as in other groups, there are naturally individuals for whom high wages for short hours are the greatest attraction a situation can offer and, in many cases, economic pressure at home prohibits any consideration other than substantial and speedy remuneration. However, the assumption that this is the guiding principle of the vast majority of school-leavers and their advisers when they first enter industry would appear, from the present data, to be unwarranted.

IV. SUMMARY

The main points discussed in this paper may be summarized as follows:

(i) The relationship between employability and the intelligence test scores proved to be stronger than that between employability and scholastic ability.

(ii) The relationships between employability and both appearance and manner, and home environment, were found to be stronger than either of the connexions referred to in (i).

(iii) Parents appeared to be on the whole more vocationally ambitious for their sons than the boys themselves.

(iv) Very little agreement was found among employers as to the most important qualities required by their juvenile employees.

(v) A comparison of employability with temperament and character traits suggested that in practice the distribution of boys among good and bad grade situations may to some extent be determined by the possession or lack of certain such traits.

(vi) Remuneration was found to play a smaller part in determining occupational satisfaction than did the type of work done or the physical conditions under which it is carried out.

V. CONCLUSION

Owing to the small number of boys so far considered and the short time during which they have been employed, every conclusion is of a tentative nature. Material on similar lines concerning the second and third groups of school-leavers is being collected and the investigator is keeping in touch with members of the first group and their employers. The examination of these further data will furnish some guide as to the value of these first suggestions.

(Manuscript received 10 March 1938)

THE EFFECT OF ADAPTATION UPON VISUAL ACUITY

BY K. J. W. CRAIK

From the Psychological Laboratory, Cambridge

- I. *Introduction* (p. 252).
- II. *Apparatus and procedure* (pp. 253-257).
 - (1) *General description of apparatus* (pp. 253-254).
 - (2) *Details of apparatus* (pp. 254-255).
 - (3) *Calibration* (p. 256).
 - (4) *Experimental procedure* (pp. 256-257).
- III. *Results* (p. 257).
- IV. *Discussion* (pp. 258-263).
- V. *Conclusions* (pp. 263-265).
- VI. *Summary* (p. 266).
- References* (p. 266).

I. INTRODUCTION

It was found by Mellone and Rawdon-Smith (1935, unpublished), and in a more extended investigation by the present author (1938), that visual adaptation has a considerable effect upon intensity discrimination. The differential threshold was found to be lowest when tested at the same illumination to which the eye was adapted, and to be raised considerably where these 'test' and 'adapt' illuminations were widely different.

It seemed that a similar investigation of visual acuity might throw some light on the relation between this and intensity discrimination. A positive result was to be anticipated from the experiments of Lythgoe (1932) in which acuity was diminished by a surrounding field, except when that field had approximately 1/10th the brightness of the smaller field containing the test object; for spatial and temporal proximity of an adapting field appear in general to have very similar effects.

II. APPARATUS AND PROCEDURE

(1) *General description of apparatus*

The requirements of the apparatus for this investigation may be summarized as follows. To secure adaptation there should be a uniformly illuminated field, preferably subtending more than 30° at the eye, and capable of variation over a wide range of brightness. The 'test' field should be approximately the same size in order that the retina may be exposed to the test illumination over a wider area than that covered by the test object. Otherwise, glare and 'contrast' effects may disturb the acuity (Lythgoe, 1932). The test field should be variable over the same range of brightness as the adapting field, but independently of it. Within the test field should appear an object with parts whose angular separation is variable from 0 to 40 min. Its black portions must be as completely dark as possible, and the ratio of the widths of the black and white portions should remain constant while the separation varies. The 'adapt' and 'test' fields must have the same spectral composition and be rapidly interchangeable without requiring movement of the subject's head or eyes.

Considerable difficulty was encountered in meeting all these requirements simultaneously, and it was decided to employ a principle which rendered possible the use of high illuminations, rapid interchange of 'test' and 'adapt' fields, and convenient measurement of widely different acuities, though it entailed a field subtending only 16° .

The light from an electric bulb, rendered nearly parallel by a condenser, filled a large lens at whose principal focus the eye was placed. This lens appeared uniformly illuminated by the 'Maxwellian view'. Close to the condenser was a test object consisting of two parallel strips of thin, blackened metal, joined at their ends so as to form a rectangular frame with a central gap, equal in width to one of the strips. This frame could be rotated through 90° about its longer axis, so that the strips and the gap appeared to become narrower, if viewed from a fixed position. A small lens of short focus, mounted between the condenser and the large lens, cast a much reduced image of the strips (and of the bright background formed by the condenser) in front of the large lens, so that the image was seen sharply focused in the centre of the large field. The small lens was moved out of the optical path in the 'adapt' position, and a double filter-frame near the eyepiece, operated by a key-arm in conjunction with the small lens, enabled the illumination to be altered.

The appearance of the 'test' field as seen by the subject is shown in Fig. 1.

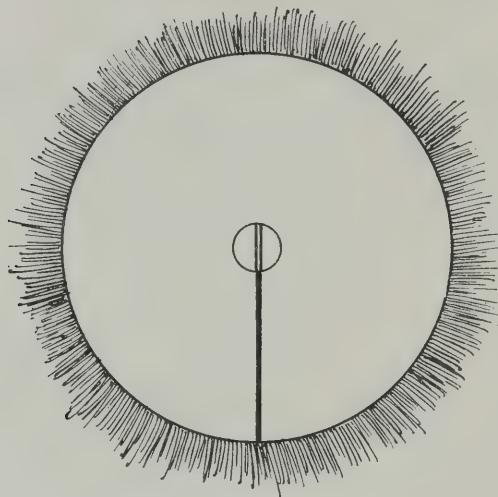


Fig. 1. Appearance of field with test object in centre.

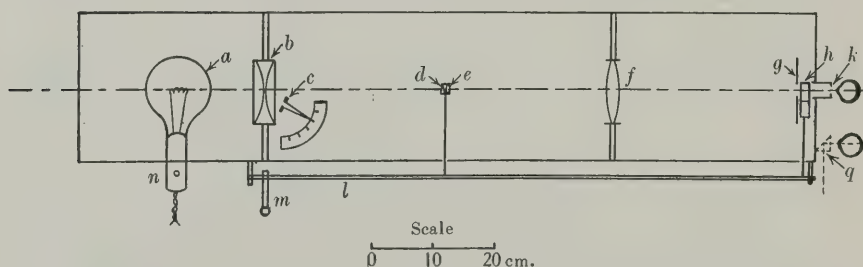


Fig. 2. Diagram of apparatus. *a*, 100 W. pearl bulb. *b*, condenser. *c*, test object and scale. *d*, prism, and *e* lens, both in movable frame. *f*, field lens. *g*, stop. *h*, double filter frame. *k*, artificial pupil. *l*, rod. *m*, key-arm. *n*, switch. *q*, prism for calibration.

(2) *Details of apparatus*

A condenser *b* (Fig. 2), 10 cm. in diameter, is placed in front of a 100 W. 200 V. 'pearl' bulb *a*. Immediately on the other side of the condenser is the test object *c*, consisting of two blackened metal strips 10 cm. $\times$ 6 mm. $\times$ 0.2 mm., fastened together, with a 6 mm. gap between them, by means of cross-strips at their ends. This rectangular frame is pivoted at one end and fitted with a pointer and scale. It can be rotated by a slow motion gear and knob; any back-lash in the latter is without effect on the calibration as the pointer is rigidly attached to the test object.

60 cm. from the condenser is a lens *f*, 10 cm. in diameter and 24 cm. in focal length,

which receives light from the condenser and, when the eye is placed at k , its principal focus, provides a uniformly illuminated field, 16° in diameter.

32 cm. from the condenser is a smaller lens e having a diameter of 10 mm. and a focal length of 10 mm., which casts a much reduced image of the test object between it and the large lens f . The test object c had to be placed slightly off the optical axis in order to prevent it from casting a shadow over the Maxwellian field of f . To restore the image of the test object to the centre of the system, a 6° prism, d , 10 mm. in diameter, was mounted in the same cell as the lens e . This cell, constructed of metal 0.2 mm. thick, was mounted on a metal arm presenting only a thin edge, 1 mm. wide, in the direction of the optical axis. These precautions were taken to avoid the presence of large black areas in the test field, since these would interfere considerably with acuity in neighbouring regions. The arm is pivoted at its base so as to enable the small lens and prism to be swung out of the way.

This small lens has two functions. First, it introduces an image of the test object into the field of view, without the mechanical inconvenience of moving the object and its scale. Secondly, it reduces the apparent size of the test object about tenfold, permitting the employment of a relatively large and robust object concurrently with short optical paths and compact design.

This image is unavoidably magnified about $2\frac{1}{2}$ times by the lens f which provides the large field, but it is still considerably smaller than the apparent size of the test object at the same distance. The condenser b has also two functions—to concentrate light on the lens f , and to provide a uniformly illuminated background for the test object. The distances of the lenses are so arranged that the bright background of the test object has very nearly the same brightness as the surrounding field, and so that the apparent distance of the test object is infinity. The apparent length of the test object is 2° , and, at maximum separation, the centres of the strips subtend $14'$ at the eye; this does not represent the lowest measurable values of acuity, which are nearer $30'$, but it enables acuity to be measured continuously from its keenest value down to 5% of this; and the further reduction to 1% of its original value is not essential for the present experiment.

A stop is fitted at g to minimize scattered light, and a double filter-frame, h , is mounted on a pivoted lever. Each half of this frame holds two Wratten neutral filters, 25 mm. sq., mounted in B glass, or alternatively one filter and a clear glass square, or a blank stop. An artificial pupil 3 mm. in diameter is fitted at k , in front of the subject's eye.

A rod l , rotated through a small angle by a key-arm m moves the filter frame so that either half is in the optical path, and at the same time swings the lens and prism (d and e) in or out of the field. Thus, in the 'adapt' position, a uniform field of 16° is exposed; on depressing the key to the 'test' position, the other set of filters becomes operative and the small lens swings to the centre of the field, exposing the test object. The links actuating the filter frame can be detached when equal test and adapt illuminations are to be employed, thus avoiding duplication of filters. The change from 'adapt' to 'test' position occupies about 0.1 sec.

A switch at n enables the light to be extinguished while the sets of filters are changed between series of readings.

The optical system is mounted inside a light-tight box having ventilating holes and light traps. The whole is mounted on an adjustable table, and fitted with a chin-rest for the subject.

(3) Calibration

The calibration falls into three portions:

(1) The angles subtended at the eye by the field, the small circle containing the test object, and the angle between the lines of the test object at maximum separation, were measured in the following way. A right-angled prism was fixed to the box at *q*, and a screen with a millimetre scale placed three metres distant. This permitted binocular comparison between the field and the scale; slight difficulties owing to the different distances of the objects were counteracted by numerous measurements. Calculation based on this method gave results in good agreement with the magnitudes computed from the distances and focal lengths of the lenses.

(2) The maximum angular separation of the test strips being thus found, the scale of *c* could be calibrated if the relative separations for different positions of the pointer on *c* were known. Although this could be calculated from the sine of the angle of rotation of the pointer, this was considered inadvisable owing to the finite thickness of the strips and the possibility of slight mechanical inaccuracies. A small lens, similar to *e*, was placed between *e* and *f*, the lens *f* being temporarily removed. This additional lens cast a greatly enlarged image of the image of the test object on a screen with a millimetre scale, placed at *g*. By means of the pointer, the test object was set to give the distances, in millimetres on this scale, corresponding to the required angles of separation, and the pointer-scale marked accordingly.

(3) The calibration of the field brightness was accomplished by a selenium-oxide photo-cell of the generator type. This was clipped over the artificial pupil and the reading of the microammeter noted; this reading was then reproduced by illuminating a magnesium-oxide screen, subtending 16° at the cell, against a background of darkness. The artificial pupil was still fixed to the photo-cell. This diffusing screen was illuminated by a 1000 c.p. projector lamp, calibrated in turn by a 20 c.p. N.P.L. sub-standard lamp. The illumination on the screen was calculated from the inverse square law. The maximum illumination of the field was found to be approximately 10,000 equivalent ft. candles (the unit used by Lythgoe (1932) etc.). Similar measurements were made with the 1/10th and 1/100th transmission filters in position; for lower illuminations the transmissions of the filters were accepted at their assigned values. The minimum intensity was 0.0001 e.f.c., obtained by two 1/10,000th transmission filters in series.

Accurate calibration of the absolute brightness of a 'Maxwellian view' is somewhat difficult, but the above method gives sufficient precision for this investigation.

(4) Experimental procedure

The subject was seated with his head fixed in a chin rest, in such a position that the field appeared to him most uniformly illuminated, and so that the beam of light from the apparatus fell symmetrically upon the pupil of his right eye, which was then exposed to the 'adapt' field for a period sufficient to ensure nearly complete adaptation to the illumination employed. This varied from 5 min. for adaptation to 'the highest in-

tensities, to 20 min. for dark adaptation. The test object, with the required illumination of the whole field, was then exposed for periods of 2 sec., its separation being increased or decreased progressively between each exposure; the subject indicated whether or not he was able to resolve the lines. Between these exposures adaptation was restored for 1 to 5 min.; the longer periods were necessary for the recovery of scotopic sensitivity after exposure to the highest illuminations. These periods, as well as those for preliminary adaptation, were found empirically to ensure a sufficiently steady state, without prolonging the experiment so far as to cause undue fatigue. No particular fixation point was provided; the subject was instructed to fixate the centre of the field and to do his best to resolve the test object in the time provided. Trials at equal 'test' and 'adapt' illuminations with longer exposures showed that 2 sec. was sufficient for reliable judgment. It appeared that foveal and slightly eccentric fixation were automatically employed at those illuminations at which each was most effective.

Two experienced subjects (R.S.S. and R.C.O.), both of whom had slight myopia, were employed in the main experiment, and two others (M.D.V. and K.J.W.C.), both emmetropes, in trials at equal test and adapt illuminations.

III. RESULTS

In Fig. 3, the acuity of these four subjects is plotted linearly against the logarithm of the illumination. In Fig. 4, log acuity is plotted against log illumination, for equal test and adapt intensities. Log acuity is given as $\log 1/\theta$, where θ is the angle in minutes subtended by the centres of the strips. This method was used by Schlaer (1937). In Fig. 5, log acuity is shown against the logarithm of the testing illumination, for different values of adapting intensity; in general it is highest where the two illuminations are equal. At 1/1000th of the adapting illumination, acuity is extremely poor—approximately 0.1—and the same at 10,000 e.f.c. unless the eye has been adapted to a higher illumination than 100 e.f.c.

The effect of large pupil size on acuity is shown in Fig. 6, in which acuity at equal 'test' and 'adapt' illuminations is shown with the normal pupil, and after dilation to 6 mm. diameter with Euphthalmine, a mydriatic which has little effect on accommodation. There is a slight decrease in acuity at the highest illuminations.

IV. DISCUSSION

First of all, some particular features of the results may be noted. The curve relating acuity to the logarithm of intensity appears to be approximately a straight line for illuminations between 0.001 and 10 e.f.c.

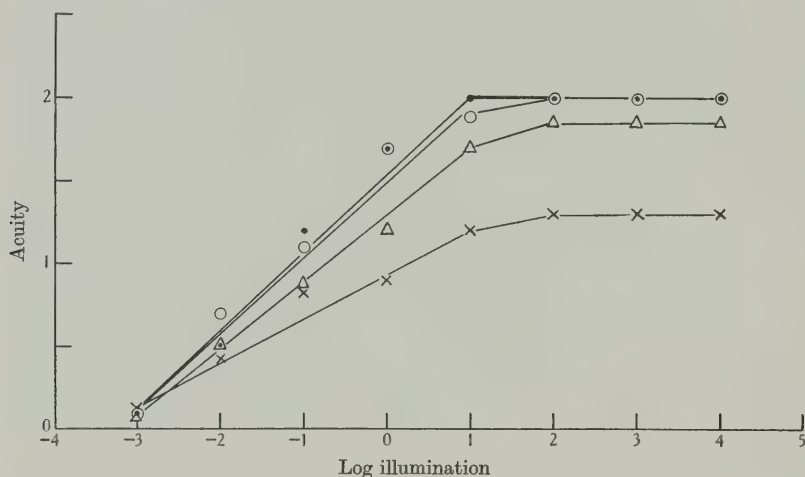


Fig. 3. Acuity against log illumination. Four subjects.

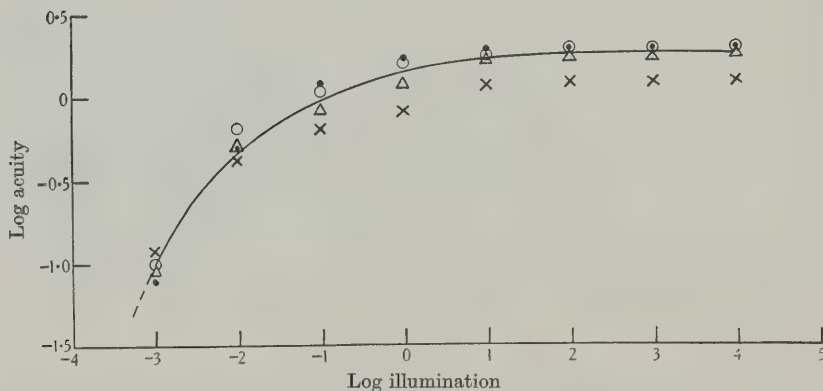


Fig. 4. Log acuity against log illumination; same data as Fig. 3.

(see Fig. 2), thus supporting Lythgoe's results (1932). From 10 to 10,000 e.f.c., the curve is flatter and seems to be quite level above 100 e.f.c. This disagrees with the data of Lythgoe (1932) and Schlaer (1937) using a C-shaped test object; it accords, however, with the results

of Banister, Hartridge and Lythgoe (1927), and of Lythgoe (1932) for a grating. Thus it seems that the discrepancy is more likely to be due to the

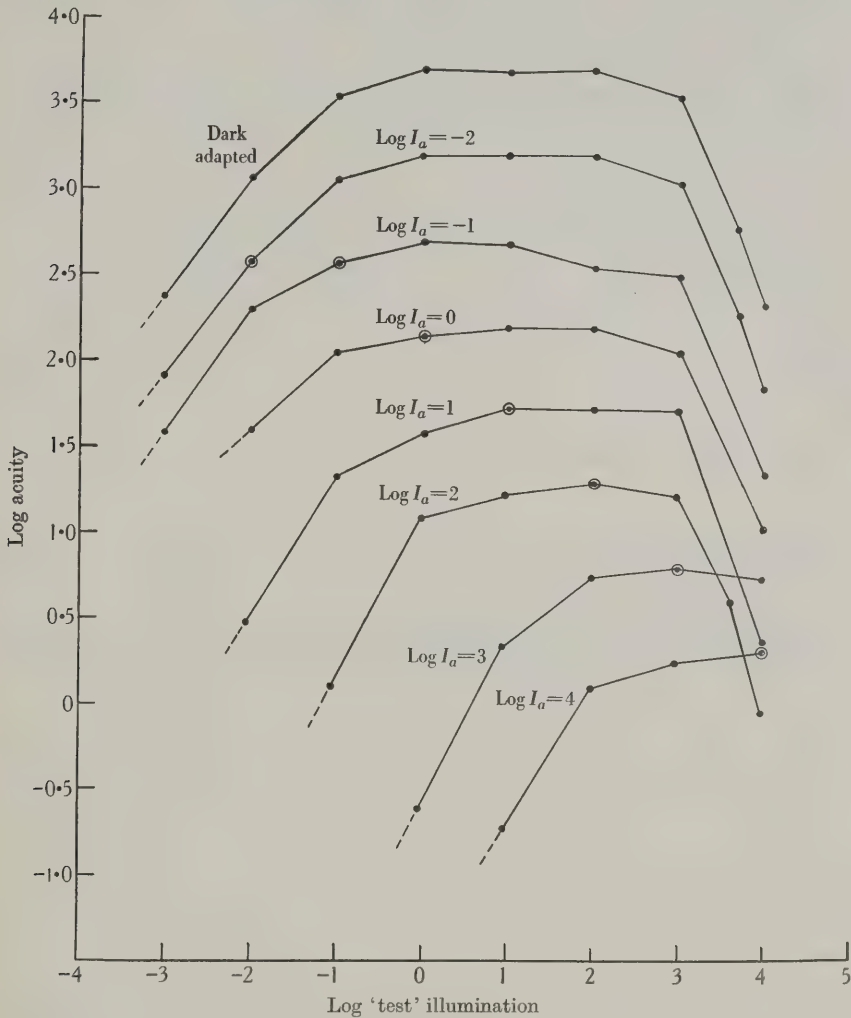


Fig. 5. Log acuity against log 'test' illumination, at various adapting intensities, marked beside the curves. Circles represent equal 'test' and 'adapt' illuminations. All curves above the lowest are shifted upwards by 0.5 on the ordinate scale, to avoid confusion.

test object employed than to any cause such as insufficient size of adapting field. Very inadequate adaptation would certainly cause such a deterioration of acuity, as the results of the present experiment confirm

by means of successive adaptation; but the agreement of these data and those of Schlaer, who employed a 30° field, suggest that the 16° field used in the present experiment was adequate, and that the curve is indeed flat for a grating or parallel-line object.

Secondly, the present results neither confirm nor disprove the existence of a photopic-scotopic inflexion, since the maximum width of the test object was insufficient to give points on the curve at the requisite very low illuminations.

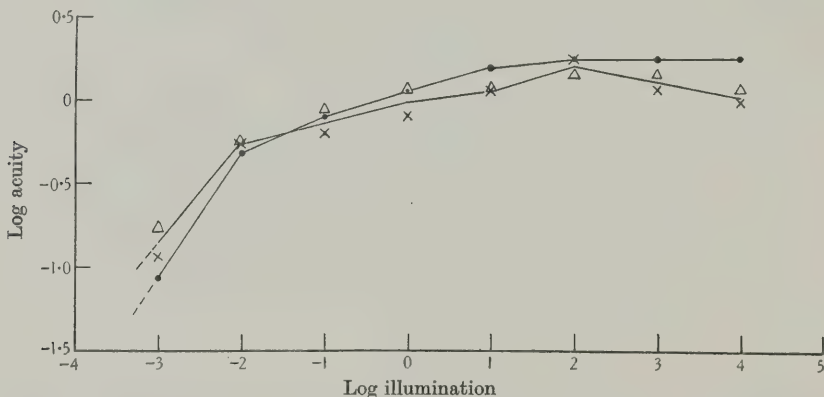


Fig. 6. Log acuity against log illumination before and after dilation of pupil with Euphthalmine. Dots, normal pupil, and artificial pupil 3 mm. diam. Crosses and triangles, two different occasions after Euphthalmine.

There are two main theories of acuity—those of Hecht (1928) and Hartridge (1918, 1923). The first endeavours to account for the effect of illumination on acuity by supposing that the sensitivity of individual rods and cones is normally distributed. Acuity will then be impaired at very low illuminations since few fresh receptors are available with the appropriate sensitivity to enable them to come into operation. This theory fits ill with the fact, demonstrated in Hartline's electrophysiological studies of nerve impulses in the vertebrate eye (1938), that the 'on fibres' give a response which increases in frequency with intensity of stimulation. Further, if the theory were correct, it might be expected that acuity would be extremely good at high illuminations after dark adaptation; for all receptors would have their maximum sensitivity, and the greatest number would therefore be available at any given illumination. The present results show, however, that under such conditions acuity is extremely bad. It is also found that if the space between the two lines of a test object which is just below the resolving power of the eye be

filled in, for a certain length, this part appears much darker than the rest of the object. This can be seen if Fig. 7 be observed from a distance of about 9 ft. (The same principle is of course used to produce greys in half-tone prints.) Now if two strips are resolved if, and only if, a stimulated receptor falls between their images, this is hard to explain; for two unresolved lines must have no stimulated receptor between them, and the presence or absence of black filling should make no apparent difference. If, on the other hand, resolution is due to discrimination of intensity gradients, the phenomenon is easily explicable.

Hartridge's theory, on the other hand, offers a most attractive explanation of the way in which acuity under optimal conditions and illumination exceeds what might be expected from the aberrations of the dioptric system of the eye; but it makes no attempt to explain the effect of varying the illumination on acuity and does not appear able to do so. For according to Hartridge's calculations, the illumination gradients on

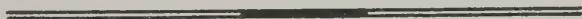


Fig. 7.

the retina which correspond to the edges of objects do not exceed 5μ or 70 sec. of arc in effective width, and no further separation of the strips beyond 5 min. of arc (from centre to centre) can increase the brightness differences on the retina. Yet it is well known that under very weak illumination separation up to 30 min. is effective in securing resolution.

Again, relative localization of two points on the retina does not appear to diminish greatly with decreasing illumination. It is possible under such conditions to provide good localization by a framework of faint fixation points without either improving or impairing acuity.

The present experiment provides four main results:

(1) Acuity is not dependent simply upon physical illumination but also upon the state of adaptation of the eye. Thus at 10,000 e.f.c. acuity may be 2 if the eye is adapted to that illumination or less than 0.07 if it is dark adapted; while at 10 e.f.c. acuity may be 1.9 if the eye is adapted to 10 e.f.c. but 0.07 if it is adapted to 10,000 e.f.c.

(2) Acuity appears to be highest when the 'test' and 'adapt' illuminations are equal, except when the adapting illumination is less than 10 e.f.c. Lythgoe (1932) found acuity to be very slightly better when the surround was 1/10th as bright as the test field, than when it was equally bright—in this case 12.6 e.f.c. This occurred almost at the illumination where a marked effect of a somewhat similar kind was found in the

present experiment, which may be explained as follows: over the range of adapting illuminations above 10 e.f.c., the highest acuity is found when the 'test' and 'adapt' illuminations are equal, and this may be taken as the general rule. At lower adapting illuminations, however, the general decline of acuity even at equal 'test' and 'adapt' intensities enters as a complicating factor. The result will be that acuity may be higher at an illumination somewhat above the adapting one; but it is usually equal or higher still when the eye is adapted to, and tested at, that higher illumination.

(3) Acuity becomes extremely low

(a) when the 'test' illumination is 100 times as high as the adapting one, if this latter is higher than 100 e.f.c., or

(b) when the 'test' illumination is 10,000 e.f.c., if the adapting one is less than 100 e.f.c.

Acuity is also very low when the 'test' intensity is 1/1000th of the adapting one, but over a wider range of adapting intensities, i.e. from 1 to 10,000 e.f.c. Let those 'test' illuminations (above and below the adapting intensity) at which acuity becomes extremely poor, e.g. 0.07, be called the upper and lower limits of acuity for that level of adaptation. Then it appears that the lower limit is proportional to the adapting illumination over a range of 4 log units, and the upper limit over a range of 2 log units only (10,000 e.f.c. being the highest adapting illumination obtainable in the apparatus). A possible explanation seems to be as follows: The progressive shift of the upper limit is probably due to adaptation of the cones; that of the lower limit to that of the rods. Perhaps the cones do not increase in sensitivity below an adapting illumination of 100 e.f.c.; and as central vision is automatically used at high testing intensities, whether the eye be previously bright or dark adapted, there would be no increase in "dazzle" nor decrease in acuity. At low test intensities, however, the eye is involuntarily moved so as to bring the image of the test object upon the retinal area most sensitive to that illumination, which contains many rods and is capable of further dark adaptation and extension of the lower limit of acuity.

(4) The present results are very similar to those obtained by the writer (1938) for the effect of adaptation upon brightness discrimination. It has been suggested above that neither Hecht's nor Hartridge's theory affords an adequate explanation of the influence of illumination on acuity, and further that the present results, showing a decrease in the acuity of the dark adapted eye at high intensities, are contrary to what would be expected on Hecht's theory.

It has been argued by Legrand (1935) that the main factor in the deterioration of acuity at low illuminations is large pupil area, which exposes the non-spherical part of the lens. Legrand claimed to have measured acuity independently of the dioptric system of the lens by means of interference fringes formed on the retina, and to have found little decrease at low illuminations. This seemed unlikely, since in most studies of acuity, including the present experiments, an artificial pupil of small size has been used; but as a further test the pupil was dilated with Euphthalmine, which increased its diameter to 6 mm. when exposed to 10,000 e.f.c. The artificial pupil was of course removed from the apparatus. A slight decrease in acuity, from 1.81 to 1.1, was found at 10,000 e.f.c. (Fig. 6), but this was not of the same order as that occurring under low illuminations or inappropriate adaptation.

V. CONCLUSIONS

Is it possible, then, to suggest any explanation based on intensity discrimination? Kravkov (1938) suggests a connexion between the two, and Schlaer applies to his acuity data the same equations which Hecht (1934) had derived for brightness discrimination, but these throw little light upon the way in which increase in the size of the test object can make such discrimination possible. On the other hand, it has been shown that increase in the size of a test patch lowers the absolute threshold, until the test patch subtends more than 20° (Reeves, 1918). The same effect has been found for the differential threshold, until the patch subtends more than 45 min. (Lasareff, 1911). Such an increase in size could, then, compensate to some extent for the effect of decreased illumination upon intensity discrimination.

It is not to be expected that the data for acuity and brightness discrimination should be identical. The relation between the width of a test area and the reciprocal of the differential threshold is not exactly linear, as may be seen from Lasareff's data. Pupil size and the form of the lens will affect the very crowded intensity gradients which may be involved in acuity more than the larger and less critical areas used for brightness discrimination. Again, however low the illumination, it is always possible to obtain a differential brightness judgment if the variable patch or intermittent stimulus is made so bright as to be above the absolute threshold, though the background be below it; whereas Lasareff's results show that there is a limit to the effective increase in size of the test area.

In Fig. 8 the classical data of König and Brodhun (1889) on intensity

discrimination, of König (1897) on acuity, and the recent results of Steinhardt (1936) on intensity discrimination and of Schlaer on acuity, are superimposed. $\log \Delta I/I$, and $\log \theta - 3.8$ (where θ is the subtended angle in minutes, taking an acuity of 1 as 1 min.) are plotted against $\log I$. A similar superposition is shown in Fig. 9, using the present results for acuity and those of the writer (1938) for brightness discrimination. In both cases, considerable similarity will be seen. This would be of little account if any clear-cut and insuperable difference could be found; the

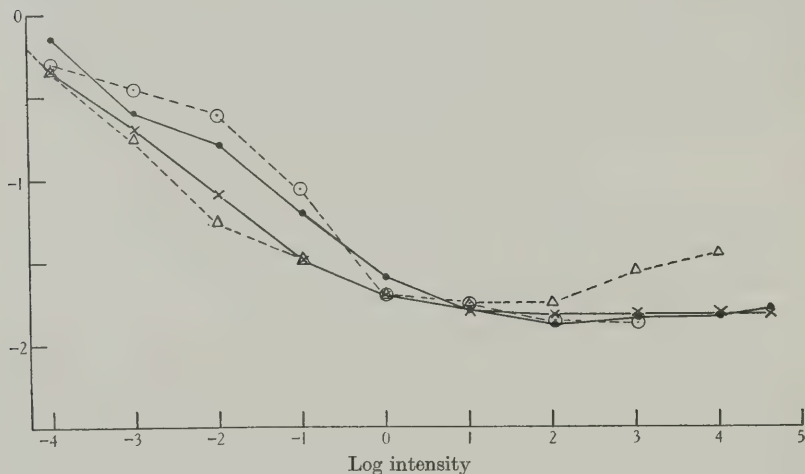


Fig. 8. Brightness discrimination data of König and Brodhun (broken line and triangles) and of Steinhardt (dots) and acuity data of König (broken line and circles) and of Schlaer (crosses), superimposed. $\log \Delta I/I$ and $\log \theta - 3.8$ are used. To avoid confusion, only points representing integral log intensities, or the mean of the two nearest points, are plotted.

supposed deterioration of brightness discrimination at high illuminations, and the continued improvement of acuity, have long been held to constitute such a difference. Recent work, such as that of Steinhardt, suggests, however, that the deterioration of brightness discrimination is not so marked as had been thought, and probably does not occur with test fields smaller than 5° . Again, the decrease in acuity at 1000 and 10,000 e.f.c. resulting from dilation of the pupil by Euphthalmine (Fig. 6) suggests that, normally, constriction of the pupil may compensate for a latent decrease in acuity consequent upon impaired brightness discrimination. A final point of similarity between the two visual functions is the need for very small eye movements; steady indirect fixation of the

acuity test object in a somewhat similar experiment was found, under very weak illuminations, to lead to its vanishing altogether.

The effect of field size upon the absolute and differential thresholds has various possible explanations; it may, for instance, be due to some kind of retinal summation (cf. Adrian and Matthews, 1927, 1928; Elsberg and Spotnitz, 1937). A simpler mechanism might involve spatial adaptation, such as the influence of a 'surround'. A change in intensity seems to be most readily appreciated when the eye is adapted to one illumina-

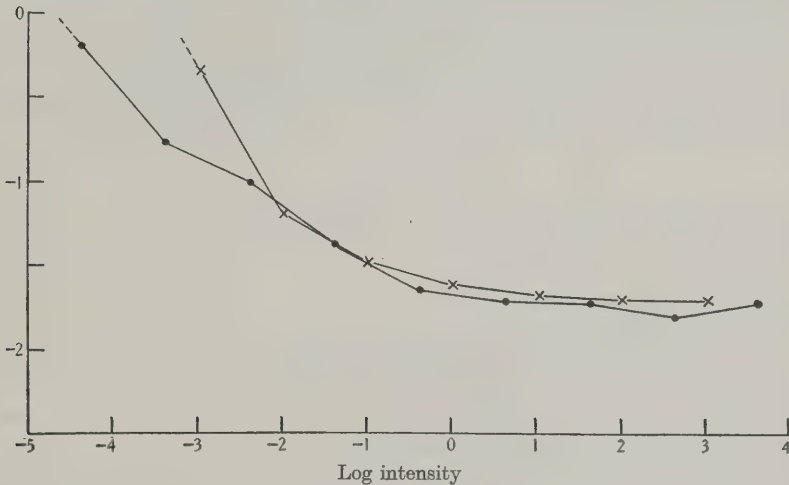


Fig. 9. The present acuity data and the writer's brightness discrimination results superimposed as in Fig. 8. Dots, $\text{Log } \Delta I/I$; crosses, $\log \theta - 3.8$.

tion and is in the initial stages of exposure to a slightly greater or smaller illumination (Helmholtz, 1866; Stiles, 1929; Hecht, 1935); while the work of Lythgoe and Tansley (1929) on surrounds shows a similar action to that of temporal adaptation. The effect of a surround is felt over several degrees, but diminishes with its distance from the test area. It follows that the centre of a large area of the retina, in darkness, will be more dark-adapted than the parts near a bright area, and will therefore give a greater response when, as a result of eye-movements, it is exposed to light. Conversely, those parts of the retina which are in the centre of a large bright area will give a greater decrease in response when the image of a dark strip falls on them. The difference in response between receptors, or in the same receptor at successive movements, will therefore be greater, within limits, when the areas of the bright and dark strips are increased.

VI. SUMMARY

Visual acuity has been tested immediately after exposure of the eye to a wide range of illuminations equal and different from those at which acuity was measured. It was found that acuity was highest (approximately 2) when 'test' and 'adapt' illuminations were equal; at illuminations far above and below the test intensity, it deteriorated greatly. The results are very similar to those previously found for brightness discrimination. A tentative explanation is proposed for these effects and for some anomalous features of them.

I am indebted to Mr R. S. Sturdy for assistance in taking readings.

REFERENCES

- ADRIAN, E. D. & MATTHEWS, R. (1927). *J. Physiol.* LXIV, 279.
 ——— (1928). *J. Physiol.* LXV, 273.
 BANISTER, H., HARTRIDGE, H. & LYTHGOE, R. J. (1927). *Brit. J. Ophthalm.* XI, 131.
 CRAIK, K. J. W. (1938). *J. Physiol.* XCII, 406.
 ELSBERG, C. A. & SPOTNITZ, H. (1937). *Bull. neurol. Inst. N.Y.* VI, 253.
 HARTLINE, H. K. (1938). *Amer. J. Physiol.* CXXI, 400.
 HARTRIDGE, H. (1918). *J. Physiol.* LII, 175.
 ——— (1923). *J. Physiol.* LVII, 52.
 HECHT, S. (1928). *J. gen. Physiol.* XI, 225.
 ——— (1934). *Handbook of Experimental Psychology*, p. 734. Ed. by C. Murchison.
 ——— (1935). *J. gen. Physiol.* XVIII, 767.
 HELMHOLTZ, H. VON (1866). *Physiological Optics*, Eng. trans, 1924, II, 179, 237–8, 266, 281.
 KÖNIG, A. (1897). *S.B. Akad. Wiss. Wien*, p. 559.
 KÖNIG, A. AND BRODHUN, E. (1889). *S.B. Akad. Wiss. Wien*, p. 641.
 KRAVCOV, S. V. (1938). *Acta ophthalm., Kbh.*, XVI, 385.
 LASAREFF, P. (1911). *Pflüg. Arch. ges. Physiol.* CXLII, 235.
 LEGRAND, Y. (1935). *C.R. Acad. Sci., Paris*, CC, 490.
 LYTHGOE, R. J. (1932). *Med. Res. Coun. Spec. Rep. Ser. No.* 173.
 LYTHGOE, R. J. & TANSLEY, K. (1929). *Proc. roy. Soc. B*, CV, 60.
 REEVES, P. (1918). *Astrophys. J.* XLVII, 143.
 SCHLAER, S. (1937). *J. gen. Physiol.* XXI, 165.
 STEINHARDT, J. (1936). *J. gen. Physiol.* XX, 185.
 STILES, W. S. (1929). *Ill. Res. Tech. Paper*, VIII.

(Manuscript received 9 May 1938)

RELIABILITY OF MENTAL TESTS

BY ROBERT W. B. JACKSON

Colonial Department, University of London Institute of Education

- I. *Introduction and statement of the psychological problems considered* (pp. 267-269).
- II. *Application of the method of testing statistical hypotheses to the solution of problems (1) and (2)* (pp. 270-277).
- III. *Estimation of the trial effect* (pp. 277-278).
- IV. *Estimation of the sensitivity of the test* (pp. 278-279).
- V. *Examples* (pp. 280-283).
- VI. *Relationship between the 'reliability' and the 'sensitivity' of a test* (pp. 284-285).
- VII. *Sensitivity of a 'whole' test and that of a 'half' test* (pp. 285-286).
- VIII. *Summary* (pp. 286-287).
- References* (p. 287).

I. INTRODUCTION AND STATEMENT OF THE PSYCHOLOGICAL PROBLEMS CONSIDERED

MENTAL tests may be used for many purposes, the diagnosis of mental deficiency, the grading of pupils, and in vocational guidance and selection. Whatever the purpose, the test is assumed to measure some capacity of the individual tested, and we use the score, or some function of the score, obtained by an individual on the test as an estimate of the individual's ability or capacity.

A test may be considered as reliable if it consistently measures whatever capacity it was designed to measure. The correlation between scores obtained by individuals on (1) duplicate forms of the same test given at different times, or (2) two applications of the same test, or (3) two parts of the same test, has been used as an estimate of the reliability, and the correlation coefficient so obtained has been termed the 'reliability coefficient' of the test.¹

This is the definition, and measure, of reliability in common use. It seems to me, however, that we may consider this problem from another viewpoint. A mental test is, or may be considered as, an instrument for measuring capacity in much the same sense, for example, as a yard-stick

¹ In the third case the reliability of the whole test is found by using the Spearman-Brown formula.

is an instrument for measuring length. It may be noticed that the appropriateness of the yard-stick in any given case depends on the relationship between the scale in which it is subdivided (e.g. inches, and tenths of an inch) and the variability of the objects measured. If these differ, for example, by not more than one-tenth of an inch, then, probably, an ordinary yard-stick would not be considered as a sufficiently accurate tool. On the other hand the yard-stick is frequently considered as satisfactory when the objects differ by as much as a foot. This idea seems to be fundamental and I shall use it below, applying it to mental tests. These will be considered as reliable or appropriate when the inaccuracy of measurement by means of these tests is not very great compared with the variability of the capacity measured. The mental test, of course, may be a poor measuring instrument, and we generally wish to know how accurately, or inaccurately, it measures the capacity it was designed to measure. To answer this question we would proceed exactly as before in obtaining the experimental results, but, as will be shown later, the correlation coefficient is not the appropriate measure to use in this case.

The score of an individual in a particular trial of a test will be composed of a measure of the capacity of the individual, a measure of the effect common to all individuals on that particular trial, and a measure of numerous other effects including errors of sampling which we cannot measure separately but must group together. We may subdivide the measure of the effect common to all individuals on a particular trial into two parts: (1) a measure of the effect common to all individuals irrespective of the particular trial under consideration and (2) a measure of the trial effect, i.e. a measure of the effect determined by the particular trial under consideration after the common effect (1) has been allowed for. It will be seen that the trial effect must be considered independently of the measurement of the individual's capacity since it may exist whether or not the test actually measures the capacity of the individual. The interpretation which may be placed on reliability depends on the purpose for which the test is to be used. For example, in the diagnosis of mental deficiency we calculate an estimate of the individual's mental age, or intelligence quotient, from the score obtained on the test. In this case we should conclude that a test is satisfactory if it measures the capacity of the individuals accurately, and if the trial effect is not significant; if the trial effect is significant our estimate of the individual's mental age, or intelligence quotient, will depend on the particular trial of the test from which the score is taken. On the other hand, in the grading of pupils we are interested mainly in the position of the individual in the classification,

and here the trial effect is unimportant if we may assume that it is independent of the particular individual tested. In general, however, we shall be interested in the actual rather than in the relative scores, and must consider the trial effect independently of the measurement of the capacity of the individual.

The present paper, therefore, is concerned with four separate problems:

- (1) To determine if there is a significant trial effect.
- (2) To determine if the mental test actually measures the capacity of the individuals tested.
- (3) To estimate the trial effect if it exists.
- (4) To obtain an estimate of the relative importance of the random errors of measurement, with respect to the true measurement of the capacity of the individual, in determining the score of the individual on the mental test.

In the following pages I have assumed that a linear relation holds between the effects enumerated above and have considered the grouped effects as error; this grouped effect could more properly be termed the residual effect but, as we are unable to distinguish between the pure error and other effects of this group, we may consider the total residual effect as error. It was realized that the relation between these effects may not be linear but, as the true relation is in any case unknown, this form was used because the analysis was simplified and I could apply directly the methods of testing statistical hypotheses, and of estimation, devised and developed by Neyman and Pearson (1,2,3).

I have assumed also that the capacity of the individual does not change in the interval elapsing between the first test and the retest. The question of possible changes of capacity has been discussed by Thouless⁽⁸⁾ who points out that variations in the experimental results from repeated tests may be due to two factors—unreliability of the test and changes in capacity of the individuals tested. The problem of distinguishing between, and measuring, the effects of these two factors could be attacked by methods similar to the ones outlined in the following pages, but this is an entirely different problem from the one which I have considered. The experimental results which have been analysed below are not affected by changes in capacity as they were obtained by dividing the scores on a single application of the test into two equivalent parts. The results in some cases may be affected, however, and the interpretation of the results obtained by applying the following method would depend on the particular case considered.

II. APPLICATION OF THE METHOD OF TESTING STATISTICAL HYPOTHESES TO THE SOLUTION OF PROBLEMS (1) AND (2)

The event, E , that we observe consists of a set of test scores, in general s test scores, say, for each of n individuals. If correlation methods are used it is necessary to restrict ourselves to the case $s=2$, but the following method is applicable for any value of s greater than unity. The arguments will be developed for the case $s=2$ since the data available have been collected in a form suitable for correlation analysis, as shown in Table I, where y_{st} represents the score of the t th individual on the s th trial of the test. I shall consider the case $s=1, 2$; $t=1, 2, \dots, n$.

Table I

Individual	Trial I	Trial II
1	y_{11}	y_{21}
2	y_{12}	y_{22}
$\vdots$	$\vdots$	$\vdots$
n	y_{1n}	y_{2n}
Totals	$\sum_t y_{1t}$	$\sum_t y_{2t}$

The data may be obtained in several ways, as indicated in § I, by the use of:

(1) Duplicate forms of the same test; i.e. two separate forms of the test are available, the first form is used and later, after a suitable period of time has elapsed, the second form is used on the same class. It is generally assumed that the two forms of the test are equivalent, in a psychological sense, and that the capacities of the individuals have not changed in the interval elapsing between the trials.

(2) Two applications of the same test; i.e. the test is used once and repeated, after a suitable period of time, on the same class. It is assumed that the capacities of the individuals have not changed in the interval, and that the practice effect, if any, is independent of the particular individual tested.

(3) A division of the scores by some method into two parts, generally by taking the totals of the scores on the odd and even items for each individual. It is assumed that the two parts are equivalent, in a psychological sense, and that the difference in difficulty of the test items on the two parts is not sufficient to invalidate the division.

It is realized⁽⁶⁾ that the psychological interpretation of the results obtained by different experimental methods may not be the same in each case, but from the statistical point of view the problem is the same. The

concepts introduced are general, their exact interpretation in particular cases is a psychological rather than a statistical problem.

Assume that the score y_{st} of the t th individual on the s th trial of a test may be represented as a sum of independent factors or components. We are interested in an effect common to all individuals on the test irrespective of the particular trial, a trial effect, an individual effect, and the residual, or error, effect.

We may express this in a mathematical form by assuming that

$$y_{st} = A + B_s + C_t + z_{st}, \quad \dots\dots(1)$$

where $s = 1, 2; t = 1, 2, \dots, n$ represents the score of the t th individual on the s th trial of the test; A is a measure of the effect common to all individuals irrespective of the particular trial under consideration; it will be considered as a constant, and is defined as the arithmetic mean of the effects for all trials; B_s is a measure of the trial effect; it will also be considered as a constant; C_t is a measure of the individual effect, i.e. a measure of some capacity of the individual measured by the test; z_{st} is the error of measuring by means of the test applied.

We assume, further, that

- (1) z_{st} is independent of C_t ,
- (2) z_{st} is normally distributed about zero with constant standard deviation, σ , say, constant for all trials and all individuals, and
- (3) in the population from which we are sampling, C_t is normally distributed about zero with the standard deviation, σ_c , independent of the trial.

The tests are generally so constructed that the individual with the least capacity can answer the easiest questions, while the individual with the greatest capacity will be unable to complete the test. Because of this type of construction the distribution of test scores is generally such that it can be adequately represented by a normal distribution, and hence the above assumptions will define a mathematical model which should adequately represent the observed events.

Finally, since we have considered A as a common effect and defined it as the arithmetic mean of the effects for all trials and all individuals, it is necessary that

$$\sum_s B_s = 0. \quad \dots\dots(2)$$

In the particular case considered, (2) becomes

$$\left. \begin{aligned} B_1 + B_2 &= 0 \\ B_1 &= -B_2 = -B, \text{ say,} \end{aligned} \right\} \quad \dots\dots(3)$$

or

where B is now the measure of the trial effect, and we may write (1) in the form

$$\begin{cases} y_{1t} = A - B + C_t + z_{1t} \\ y_{2t} = A + B + C_t + z_{2t} \end{cases} \dots\dots(4)$$

Therefore, the score of the t th individual on the s th trial of the test is assumed to be composed of a constant term $A \pm B$, plus a measure of the capacity of the individual, C_t , and a random term z_{st} . If σ is large in comparison with σ_c , it is easily seen that we may frequently get a deviation from the true capacity, C_t , as large or larger than σ_c units due to

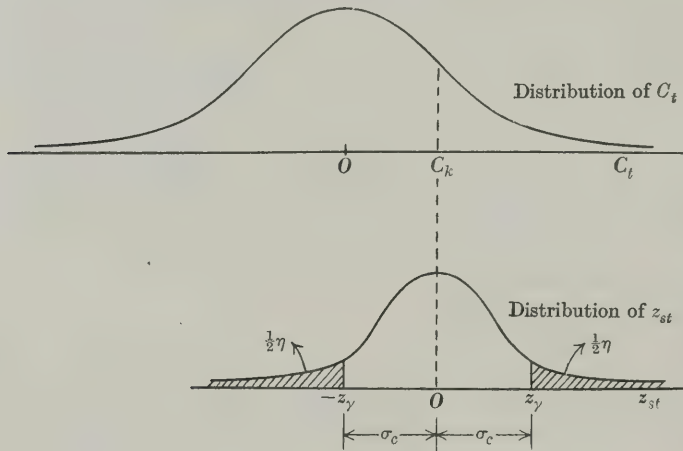


Fig. 1. Illustrating the possible effect of the z_{st} term in the score y_{st} of an individual.

chance alone if we use the score on the test as an estimate of the capacity of the individual. On the other hand, if σ is small in comparison with σ_c , then we would seldom make an error by chance alone as great or greater than σ_c units in using the score as an estimate of the capacity of the individual tested. This effect is, of course, quite different from, and independent of, the bias introduced by the trial effect as considered above.

The chance deviations may be very great in any particular case and we cannot control these, but for a particular value of γ , where γ is defined as σ_c/σ , we can determine the probability η , say, of making an error greater than or equal to σ_c units due to chance alone in using the score as an estimate of the capacity of the individual. I will describe γ as the sensitivity¹ of the mental test. It will be seen that γ is independent of the scale used in marking the test. The position is illustrated in Fig. 1.

¹ The relationship between the 'reliability' and the 'sensitivity' of a test is discussed below, see p. 284.

It is seen that the problem reduces to one of finding a value of η such that

$$P\{-z_\gamma \leq z_{st} \leq z_\gamma\} = 1 - \eta, \quad \dots\dots(5)$$

where $z_\gamma = \sigma_c$. Since z_{st} is assumed to be normally distributed about zero with constant standard deviation, σ , this reduces to finding η such that

$$\frac{1}{\sqrt{(2\pi)}} \int_{-z_\gamma/\sigma}^{z_\gamma/\sigma} e^{-\frac{1}{2}x^2} dx = 1 - \eta, \quad \dots\dots(6)$$

But $z_\gamma/\sigma = \sigma_c/\sigma = \gamma$, and hence we may obtain η directly from the tables of the normal integral for any γ . And, conversely, if we have a particular value of η , say one that we have chosen in advance, we can obtain γ from the tables of the normal integral. For example, if $\gamma = 1.96$, then $z_\gamma/\sigma = 1.96$, and from the tables we find $\eta = 0.05$. This means that in using the scores as estimates of capacity we would expect to make an error as great or greater than σ_c units by chance alone only once in twenty times. The arguments are similar for any other value of γ or η .

In the first problem we wish to determine if there is a significant trial effect. This is equivalent to testing the hypothesis that the trial effect is zero; the corresponding statistical hypothesis, H_1 , may be written in the form

$$H_1 : B = 0. \quad \dots\dots(7)$$

In the second problem we wish to determine if the mental test actually measures the capacity of the individuals tested. This is equivalent to testing the hypothesis that the sensitivity of the test is zero; the corresponding statistical hypothesis, H_2 , may be written in the form

$$H_2 : \gamma = 0. \quad \dots\dots(8)$$

The mathematical development of the tests to be used in testing these statistical hypotheses, and of the method of constructing the confidence intervals to be used in estimating B or γ if these hypotheses are rejected, will be published later⁽⁹⁾. The tests to be used, and the method of constructing the confidence intervals, will be stated here and their use explained by means of examples. This division has been made because it was thought that practical psychologists and educationists would be interested mainly in the method and results, but not so much in the mathematical theory underlying these. On the other hand this theory may be of more general interest. There are probably many other examples in psychology, and possibly in other fields, where the same, or a similar, general method could be used. The next few paragraphs will be devoted

to a brief description of the theory of testing statistical hypotheses and of estimation introduced and developed by Neyman and Pearson^(1, 2, 3).

Denote by W the same space, the co-ordinates of which are the possible results of experiment, and by w any region in this sample space. For example, in the case of mental tests the scores would be considered as the co-ordinates defining the sample space, and the observed scores on any particular experiment would determine a point, E , called the sample point, in the space of all such possible results. Denote by $P\{E \in w\}$ the probability that E , as determined by some experimental results, will fall within the region w , and by $P\{E \in w/H\}$ the probability of E falling within w as determined by some hypothesis, H .

Any test of a statistical hypothesis, H_0 , reduces to a rule of rejecting H_0 when the sample point, E , falls within some specified region, w_0 , called critical, and of refraining from doing so (i.e. 'accepting H_0 ') in other cases. The choice of critical regions is made having regard to probabilities of errors which may arise in testing hypotheses. There are two kinds of errors: (1) we may reject the hypothesis tested, H_0 , when it is true. The probability of the first kind of error is thus the probability, as determined by H_0 , of E falling within the critical region, $P\{E \in w_0/H_0\}$. (2) We may fail to reject H_0 when, in fact, it is wrong, and therefore when some other hypothesis, H' say, contradictory to H_0 is correct. If A_0 denote the region $W - w_0$, then the probability of the second kind of error will be $P\{E \in A_0/H'\}$. The region A_0 is called the region of acceptance. The critical regions used in the tests of the statistical hypotheses in the following sections are the best possible in the sense that (1) the probability of errors of the first kind is controlled at a fixed level, and (2) the probability of errors of the second kind is reduced to as low a level as possible for the class of hypotheses in which we are interested. It will be noticed that the tests of these hypotheses are of the form of the tests used in the Analysis of Variance (R. A. Fisher⁽⁵⁾) method of analysing experimental results.

The method of estimating an unknown parameter θ , say (e.g. B or γ in our case), consists essentially in determining two functions $\underline{\theta}(E)$ and $\bar{\theta}(E)$, having the property that the probability of $\underline{\theta}(E)$ falling short of the true value of θ , say θ_0 , and at the same time of $\bar{\theta}(E)$ exceeding θ_0 is exactly equal to α , a number chosen in advance and as close to unity as desired. The number α is called the 'confidence coefficient', and the interval extending from $\underline{\theta}(E)$ to $\bar{\theta}(E)$ is called the 'confidence interval'. If the functions $\underline{\theta}(E)$ and $\bar{\theta}(E)$ are determined then the process of estimation of θ consists in (1) obtaining the experimental results determining

the sample point, E , (2) calculating the values of $\underline{\theta}(E)$ and $\bar{\theta}(E)$ and (3) stating then that $\underline{\theta}(E) \leq \theta \leq \bar{\theta}(E)$. The justification for this procedure is that the probability of the statement being correct is equal to α —the value chosen in advance.

Neyman has shown that in general there is an infinity of functions $\underline{\theta}(E)$ and $\bar{\theta}(E)$ satisfying the above conditions, and therefore an infinity of systems of confidence intervals⁽³⁾. The particular system of confidence intervals used below possesses the following properties: (1) the confidence intervals cover the true value of θ , say θ_0 , in the proportion α of all such possible cases, and (2) if θ' is not the true value of θ , the confidence intervals cover θ' as rarely as possible. (The system used is known as a 'short unbiased' system of confidence intervals).

Table II

Individual	Trial I	Trial II	Totals	Squares of totals	Differences	Squares of differences
1	y_{11}	y_{21}	$y_{11} + y_{21}$	$(y_{11} + y_{21})^2$	$y_{21} - y_{11}$	$(y_{21} - y_{11})^2$
2	y_{12}	y_{22}				
$\vdots$	$\vdots$	$\vdots$				
n	y_{1n}	y_{2n}				
Totals	$\sum_t y_{1t}$	$\sum_t y_{2t}$	$\sum_t (y_{1t} + y_{2t})$	$\sum_t (y_{1t} + y_{2t})^2$	$\sum_t (y_{2t} - y_{1t})$	$\sum_t (y_{2t} - y_{1t})^2$

Table III

Statistic	Definition
$\bar{v}$	$\frac{1}{n} \sum_t (y_{2t} - y_{1t})$
S_u^2	$\frac{1}{n} \left\{ \sum_t (y_{1t} + y_{2t})^2 - \frac{(\sum_t y_{1t} + \sum_t y_{2t})^2}{n} \right\}$
S_v^2	$\frac{1}{n} \left\{ \sum_t (y_{2t} - y_{1t})^2 - \frac{(\sum_t y_{2t} - \sum_t y_{1t})^2}{n} \right\}$
χ_v^2	$\frac{1}{n} \sum_t (y_{2t} - y_{1t})^2$

The statistics to be used in testing the hypotheses H_1 and H_2 , and later in the estimation of the trial effect and the sensitivity of the test if we reject these hypotheses, may be calculated as shown in Table II. From the totals of Table II we calculate those in Table III. The calculations may be checked by use of the following identity:

$$\sum_t y_{1t}^2 + \sum_t y_{2t}^2 - \frac{1}{2n} \left\{ \left(\sum_t y_{1t} + \sum_t y_{2t} \right)^2 \right\} \equiv \frac{1}{2} n \{ (\bar{v})^2 + S_u^2 + S_v^2 \}. \dots (9)$$

(A) *Test of the hypothesis $H_1: B=0$*

(i) From Table III calculate

$$z_1 = \frac{1}{2} \log_e \left\{ \frac{(\bar{v})^2}{S_v^2/(n-1)} \right\} \quad \dots\dots(10)$$

and refer to Fisher's tables of z for $f_1=1$, and $f_2=n-1$,(ii) reject the hypothesis H_1 if z_1 is greater than the 5% (or 1%) point given in the tables,(iii) accept the hypothesis H_1 if z_1 is less than the 5% (or 1%) point given in the tables.(B) *Test of the hypothesis $H_2: \gamma=0$*

(i) From Table III calculate

$$z_2 = \frac{1}{2} \log_e \left\{ \frac{S_u^2}{S_v^2} \right\}, \quad \dots\dots(11)$$

and refer to Fisher's tables of z for $f_1=f_2=n-1$,(ii) reject the hypothesis H_2 if z_2 is greater than the 5% (or 1%) point given in the tables,(iii) accept the hypothesis H_2 if z_2 is less than the 5% (or 1%) point given in the tables.

If we know from past experience, and from the test of H_1 for the observed data, that we consider that the trial effect is zero then we may test the hypothesis H_2 under these conditions. The test is, of course, different from that considered above (B), and should be used only when the two conditions are satisfied.

(C) *Test of the hypothesis $H_2: \gamma=0$, when B is known to be zero*

(i) From Table III calculate

$$z_2' = \frac{1}{2} \log_e \left\{ \frac{S_u^2/(n-1)}{\chi_v^2/n} \right\} \quad \dots\dots(12)$$

and refer to Fisher's tables of z for $f_1=n-1$, and $f_2=n$,(ii) reject the hypothesis H_2 if z_2' is greater than the 5% (or 1%) point given in the tables,(iii) accept the hypothesis H_2 if z_2' is less than the 5% (or 1%) point given in the tables.

The question of whether we should use the 5% or 1% point is a personal one, and depends upon what risk of rejecting the hypothesis

tested when it is true (i.e. the probability of the first kind of errors) we consider as negligible or permissible. The general practice seems to be to reject the hypothesis tested if the value calculated from the sample is above the 1% point, to remain in doubt if it is between the 5% and 1% points, and to accept the hypothesis tested if it is below the 5% point.

III. ESTIMATION OF THE TRIAL EFFECT

If we accept the hypothesis $H_1: B=0$, then we may conclude that the trial effect is not significant. However, if we reject H_1 then we conclude that the trial effect is significant, and the problem arises of estimating this trial effect, B .

As shown above, this problem reduces to that of finding the upper and lower bounds of the confidence interval, $\underline{B}$ and $\bar{B}$ respectively, such that the probability of the statement

$$\underline{B} \leq B \leq \bar{B} \quad \text{.....(13)}$$

being correct is equal to $1-\epsilon$, where ϵ is any small number chosen in advance. In making the statement concerning B in the form (13) we know that we shall be correct in the proportion $1-\epsilon$ of all cases. We find that

$$\begin{aligned} \underline{B} &= \frac{1}{2} \left(\bar{v} - S_v \sqrt{\frac{l}{1-l}} \right), \\ \bar{B} &= \frac{1}{2} \left(\bar{v} + S_v \sqrt{\frac{l}{1-l}} \right), \end{aligned} \quad \text{.....(14)}$$

where l is chosen such that

$$\frac{\Gamma_{\frac{1}{2}n}}{\sqrt{\pi} \Gamma_{\frac{1}{2}(n-1)}} \int_0^l z^{-\frac{1}{2}} (1-z)^{\frac{1}{2}(n-3)} dz = 1-\epsilon. \quad \text{.....(15)}$$

We may obtain l by interpolating in Pearson's Tables of the Incomplete Beta-Function for any n and ϵ (4). The choice of a particular value of ϵ to use is a personal one, but generally $\epsilon=0.05$, or $\epsilon=0.02$, is used.

If n is larger than 101 then these tables could not be used and the first column of a small table ((7) Table V, p. 92) published by Johnson and Neyman could be recommended. This gives directly the values of $1-l$, which the authors denote by z_α , corresponding to several large values of n and to $\alpha=0.05$ and $\alpha=0.01$. The table is so arranged as to allow an easy interpolation for values of n not included in the table.

If $n \leq 31$ the alternative method of calculating B and $\bar{B}$ consists in using Fisher's tables of $t(5)$. The equations for B and $\bar{B}$ are

$$\begin{aligned} B &= \frac{1}{2} \left(\bar{v} - \frac{S_v}{\sqrt{(n-1)}} t_\epsilon \right), \\ \bar{B} &= \frac{1}{2} \left(\bar{v} + \frac{S_v}{\sqrt{(n-1)}} t_\epsilon \right), \end{aligned} \quad \dots\dots(16)$$

where t_ϵ is found by entering this table with degrees of freedom equal to $n-1$, and $P=\epsilon$.

IV. ESTIMATION OF THE SENSITIVITY OF THE TEST

If we accept the hypothesis $H_2: \gamma=0$, then we may conclude that the mental test is not sensitive enough to measure the capacities of the individuals tested and hence it would not be a satisfactory test. If we reject H_2 we conclude that the test is sensitive enough to measure the capacities of the individuals tested. How sensitive a test must be before it can be considered as satisfactory is, of course, a personal question, and could only be answered objectively by analysing the scores from some 'standard' test which has been generally accepted as giving 'satisfactory' results. This question of fixing a lower limit for the sensitivity is not a purely statistical problem. However, if we reject H_2 we would generally wish to know how sensitive the test is, or, in other words, to estimate γ , the sensitivity of the test.

As in the case of the test of the hypothesis H_2 , this problem must be divided into two parts:

- (A) Estimation of γ if the value of B is unknown, and
- (B) Estimation of γ if the value of B is known to be zero.

We are justified in using (B) only if we know from past experience, and from the test of the hypothesis H_1 for the observed data, that the trial effect is not significant. Otherwise case (A) is the appropriate one to use.

(A) Estimation of γ if the value of B is unknown

The confidence interval to be used is denoted by

$$\underline{\gamma} = + \sqrt{\frac{\theta-1}{2}} \leq \gamma \leq + \sqrt{\frac{\bar{\theta}-1}{2}} = \bar{\gamma}, \quad \dots\dots(17)$$

where

$$\begin{aligned} \underline{\theta} &= \frac{S_u^2}{S_v^2} \left(\frac{k_1}{1-k_1} \right), \\ \bar{\theta} &= \frac{S_u^2}{S_v^2} \left(\frac{1-k_1}{k_1} \right), \end{aligned} \quad \dots\dots(18)$$

and k_1 is found from the equation

$$\frac{\Gamma_{\frac{1}{2}(n-1)}}{(\Gamma_{\frac{1}{2}(n-1)})^2} \int_0^{k_1} z^{\frac{1}{2}(n-3)} (1-z)^{\frac{1}{2}(n-3)} dz = \frac{1}{2}\epsilon \quad \dots\dots(19)$$

and the tables of the Incomplete Beta-Function. Or, alternatively, by using the relation

$$\frac{1}{2} \log_e \left(\frac{1-k_1}{k_1} \right) = z_{\frac{1}{2}\epsilon} \quad \dots\dots(20)$$

finding $z_{\frac{1}{2}\epsilon}$ from Fisher's tables of z for $f_1=f_2=n-1$, and solving (20) for k_1 . For values of $n > 101$, k_1 can be determined only from Fisher's tables. The property of the confidence interval is again that the probability of the statement in the form $\underline{\gamma} \leq \gamma \leq \bar{\gamma}$ being correct is equal to the number $1-\epsilon$ fixed in advance.

(B) *Estimation of γ is B is known to be zero*

The confidence interval is defined as in (17), but $\underline{\theta}$ and $\bar{\theta}$ are now defined by

$$\left. \begin{aligned} \underline{\theta} &= \left(\frac{1-m_2}{m_2} \right) \frac{S_u^2}{\chi_v^2}, \\ \bar{\theta} &= \left(\frac{1-m_1}{m_1} \right) \frac{S_u^2}{\chi_v^2}, \end{aligned} \right\} \quad \dots\dots(21)$$

where m_1 and m_2 are defined by the equations

$$(m_2)^{\frac{1}{2}(n-1)} (1-m_2)^{\frac{1}{2}n} = (m_1)^{\frac{1}{2}(n-1)} (1-m_1)^{\frac{1}{2}n}, \quad \dots\dots(22)$$

$$\frac{\Gamma_{\frac{1}{2}(2n-1)}}{\Gamma_{\frac{1}{2}n} \Gamma_{\frac{1}{2}(n-1)}} \int_{m_1}^{m_2} y^{\frac{1}{2}(n-3)} (1-y)^{\frac{1}{2}(n-2)} dy = 1-\epsilon. \quad \dots\dots(23)$$

For values of $n > 20$ we may determine m_1 and m_2 from the equation

$$\begin{aligned} &\frac{\Gamma_{\frac{1}{2}(2n-1)}}{\Gamma_{\frac{1}{2}n} \Gamma_{\frac{1}{2}(n-1)}} \int_0^{m_1} y^{\frac{1}{2}(n-3)} (1-y)^{\frac{1}{2}(n-2)} dy \\ &= \frac{\Gamma_{\frac{1}{2}(2n-1)}}{\Gamma_{\frac{1}{2}n} \Gamma_{\frac{1}{2}(n-1)}} \int_{m_2}^1 y^{\frac{1}{2}(n-3)} (1-y)^{\frac{1}{2}(n-2)} dy = \frac{1}{2}\epsilon. \quad \dots\dots(24) \end{aligned}$$

Or, alternatively, for values of $n > 20$ we may obtain the values of m_1 and m_2 by using the following relations and Fisher's tables of z :

$$\frac{1}{2} \log_e \left\{ \left(\frac{n}{n-1} \right) \left(\frac{m_2}{1-m_2} \right) \right\} = z_{\frac{1}{2}\epsilon}, \quad \dots\dots(25)$$

$$\text{and} \quad \frac{1}{2} \log_e \left\{ \left(\frac{n-1}{n} \right) \left(\frac{1-m_1}{m_1} \right) \right\} = z'_{\frac{1}{2}\epsilon}, \quad \dots\dots(26)$$

where $z_{\frac{1}{2}\epsilon}$ is obtained from the tables for $f_1=n-1$ and $f_2=n$, and $z'_{\frac{1}{2}\epsilon}$ from the same tables for $f_1=n$ and $f_2=n-1$. For values of $n > 100$ we can obtain these values only from Fisher's tables of z .

V. EXAMPLES

The data for the following examples have been very kindly supplied by Dr Palmer O. Johnson of the University of Minnesota, U.S.A. They refer to a set of Comprehensive Examinations in Biological Sciences, prepared by Dr Johnson, and are the results for a class of 132 students at the University of Minnesota in 1935. I wish to take this opportunity to thank him for so kindly supplying these data. The original test scores are given in Table IV; it will be seen that the method used to obtain the experimental results for testing reliability is the third of those considered in Sections (1) and (2), i.e. by taking the totals of the scores on the odd and even test items for each individual. The set of tests is given in three parts: Part I, Acquisition of Vocabulary, Part II, Acquisition of Principles, and Part III, Application of Principles in Problems, or, more fully, The Ability to Apply Principles in the Solution of Biological Problems.

Table IV. *Original test scores*

Ind.	Part I		Part II		Part III		Ind.	Part I		Part II		Part III	
	Odd	Even	Odd	Even	Odd	Even		Odd	Even	Odd	Even	Odd	Even
1	46	52	63	57	34	35	32	43	39	73	64	41	40
2	38	38	71	67	42	36	33	58	54	59	67	43	37
3	64	63	70	74	41	44	34	54	50	91	97	48	54
4	73	65	119	114	50	52	35	59	59	80	83	52	59
5	61	58	109	111	57	57	36	68	65	105	126	59	62
6	34	33	75	80	30	41	37	51	57	97	96	48	53
7	57	49	88	87	33	36	38	44	37	77	72	39	36
8	66	63	83	100	55	49	39	77	76	124	116	52	57
9	25	24	68	56	20	14	40	46	57	68	70	34	41
10	30	26	59	57	35	28	41	56	53	101	98	49	45
11	45	33	55	60	43	47	42	44	46	81	67	34	32
12	73	71	106	113	47	56	43	54	46	69	65	44	43
13	45	48	56	56	35	45	44	42	51	73	81	40	44
14	55	63	81	79	51	59	45	28	30	63	74	34	29
15	66	70	102	110	48	51	46	38	39	83	89	44	42
16	49	46	94	95	43	44	47	56	52	80	80	52	54
17	64	65	97	95	40	47	48	50	47	89	77	49	50
18	45	46	84	86	39	39	49	62	54	98	120	44	47
19	61	62	91	100	51	51	50	33	42	73	76	35	39
20	52	46	85	79	41	36	51	27	28	65	71	30	32
21	67	68	106	124	49	56	52	40	37	62	77	30	30
22	59	53	86	107	49	51	53	74	73	114	113	54	61
23	55	55	104	98	41	35	54	64	54	105	96	39	46
24	51	52	78	83	40	35	55	48	51	88	81	35	46
25	50	48	91	99	51	41	56	53	46	78	84	49	52
26	52	53	82	85	43	41	57	56	52	69	71	51	53
27	28	26	64	69	34	41	58	35	41	67	70	36	46
28	42	44	55	55	38	38	59	55	45	79	80	29	41
29	59	55	87	99	40	54	60	50	46	80	95	38	41
30	24	23	50	48	30	22	61	29	28	47	74	36	34
31	42	44	75	103	46	50	62	55	56	68	65	42	49

Table IV (*continued*)

Ind.	Part I		Part II		Part III		Ind.	Part I		Part II		Part III	
	Odd	Even	Odd	Even	Odd	Even		Odd	Even	Odd	Even	Odd	Even
63	54	61	93	109	44	54	98	47	51	83	93	40	46
64	52	46	78	89	37	45	99	49	47	98	87	37	45
65	44	43	51	45	34	30	100	30	32	61	65	18	28
66	62	56	92	101	46	46	101	45	50	80	73	39	46
67	55	56	76	80	36	36	102	67	63	70	77	40	46
68	56	62	105	102	57	60	103	15	28	60	65	30	30
69	49	47	55	64	32	39	104	30	28	66	56	34	39
70	56	61	86	114	50	56	105	44	41	71	75	31	34
71	37	45	71	56	30	19	106	45	68	85	85	46	39
72	49	44	70	68	31	43	107	24	21	43	55	26	30
73	37	37	68	84	28	23	108	51	54	65	56	32	32
74	41	42	81	88	39	35	109	42	42	53	50	35	37
75	47	47	81	77	48	47	110	47	45	88	92	45	47
76	55	51	65	71	39	45	111	38	43	68	63	41	42
77	59	62	104	106	49	47	112	56	49	93	96	46	49
78	53	49	88	95	43	43	113	50	52	91	89	47	50
79	41	42	78	77	32	38	114	63	66	101	114	56	54
80	48	52	84	87	40	41	115	45	39	94	103	40	51
81	62	51	92	82	47	49	116	56	53	91	103	41	41
82	41	40	84	89	35	46	117	57	55	73	83	33	35
83	50	40	78	89	48	52	118	64	66	99	98	47	48
84	40	47	66	74	25	33	119	59	67	99	113	45	54
85	68	64	94	100	53	60	120	35	29	66	59	40	43
86	43	45	52	51	39	35	121	53	41	78	73	40	36
87	49	47	61	76	38	36	122	44	52	56	54	37	44
88	65	62	96	97	43	49	123	54	57	93	86	48	49
89	50	51	53	60	27	39	124	48	47	85	65	38	33
90	31	30	49	70	22	34	125	47	36	58	61	36	38
91	55	51	90	98	49	35	126	52	45	92	101	43	43
92	45	40	69	65	31	33	127	45	42	75	65	31	33
93	40	46	52	51	37	37	128	23	25	66	75	27	33
94	38	42	40	54	41	31	129	43	43	69	75	44	36
95	55	55	82	76	40	45	130	73	66	111	122	50	55
96	43	44	90	80	37	33	131	32	29	73	64	35	31
97	46	56	108	113	54	60	132	44	45	73	70	41	42

I analysed these data as suggested in Tables II and III, and obtained the results shown in Table V.

Table V

Statistic	Test		
	Part I	Part II	Part III
$\bar{v}$	-0.5758	2.9621	2.0682
S_v^2	522.9984	1211.4126	272.6996
$S_v'^2$	28.6685	81.5818	30.3514
χ_v^2	29.0000	—	—

Test of the hypothesis $H_1: B=0$. Substituting the values of $\bar{v}$ and S_v^2 from Table V in equation (10) we obtain the results shown in Table VI, where z_1 (5%) and z_1 (1%) are the 5 and 1% points of the distribution of z , respectively, obtained from Fisher's tables of z for $f_1=1$, and

$$f_2 = n - 1 = 131.$$

As will be seen from the last column of Table VI, the trial effect may be considered to be zero for Part I, but is different from zero for Parts II and III.

Table VI

Test	z_1	z_1 (5%)	z_1 (1%)	Result
Part I	0.2076	0.6821	0.9609	Accept H_1
Part II	1.3227	0.6821	0.9609	Reject H_1
Part III	1.4578	0.6821	0.9609	Reject H_1

Test of the hypothesis $H_2: \theta=1$, or $\gamma=0$. Substituting the values of S_u^2 and S_v^2 from Table V in equation (11) we obtain the results shown in Table VII, where z_2 (5%) and z_2 (1%) are the 5 and 1% points of the distribution of z , respectively, obtained from Fisher's tables of z for $f_1=f_2=n-1=131$. As will be seen from the last column of Table VII, we should reject H_2 in each case, and hence we should conclude that these tests are sensitive enough to measure the capacities of the individuals tested.

Table VII

Test	z_2	z_2 (5%)	z_2 (1%)	Result
Part I	1.4519	0.1428	0.2025	Reject H_2
Part II	1.3490	0.1428	0.2025	...
Part III	1.0978	0.1428	0.2025	...

It may be argued that I should have tested H_2 for Part I by the test given under (C) above, but I am not certain that the trial effect will always be zero for this test, and in any case I would have reached the same conclusion and rejected H_2 .

Estimation of the trial effect. Since I have accepted H_1 for Part I the question of estimating B , the trial effect, does not arise for this case. The confidence intervals for B for the Part II and Part III tests are given in Table VIII. I used the confidence coefficient $1-\epsilon=0.98$.

Table VIII

Test	Confidence interval for B
Part II	$0.549 \leq B \leq 2.414$
Part III	$0.467 \leq B \leq 1.601$

Estimation of γ , the sensitivity of the test. Since I am not certain that B will always be zero for Part I, I have used case (A) (p. 278) for all these tests. The confidence coefficient used is $1-\epsilon=0.98$, and the values were calculated using Fisher's tables of z . The results are shown in Table IX.

These values of γ are rather small; it would seem that the random errors of measurement, as I have termed the residual effect, may be very important in determining the actual score of an individual on the test.

Table IX

Test	Confidence interval for γ
Part I	$2.36 \leq \gamma \leq 3.63$
Part II	$2.11 \leq \gamma \leq 3.26$
Part III	$1.58 \leq \gamma \leq 2.50$

I calculated the correlation coefficient for Part I and found $r=0.90$; this test would thus be considered as satisfactory if we used the usual methods. It seems probable, therefore, that many of the mental tests in common use are not satisfactory instruments for measuring capacity. However, the value of S_u^2/S_v^2 does not depend on the value of r alone, although if the assumptions which I have made are satisfied S_u^2/S_v^2 will be of the order of $(1+r)/(1-r)$, and it would be necessary to calculate the confidence intervals for γ from the formulae given above, rather than to attempt to estimate them from the value of r .

I have assumed that the test scores are normally distributed with the same constant standard deviation on each trial of a particular test. We may test the hypothesis that the test scores are normally distributed on each trial by calculating the estimates of β_1 and β_2 , and testing whether these differ significantly from zero and three respectively (see *Tables for Statisticians and Biometricians*, Part II, Table XXXVII bis). The hypothesis that the standard deviation is the same for each trial of a test may be tested by calculating

$$z = \frac{1}{2} \log_e \left\{ \frac{S^2 y_{1t}}{S^2 y_{2t}} \right\}; f_1 = f_2 = n - 1, \quad \dots\dots(27)$$

and rejecting the hypothesis if z is significantly greater than zero (Fisher's z -test). The results for the three mental tests are given in Table X. We may accept the hypothesis that the test scores are normally distributed in each trial for five of the six cases, the sixth case—for the even scores on Part II—is doubtful. We may accept the hypothesis that the standard deviations are equal for the two trials in all three cases. These results would indicate, in this case at least, that the mathematical model which I have used adequately represents the observed events.

Table X. *Tests of the hypotheses* (1) $\beta_1=0$, (2) $\beta_2=3$ and
(3) $\sigma_{\text{odd}}=\sigma_{\text{even}}$

Test	Trial	β_1	β_1 (5%)	β_1 (1%)	β_2	β_2 (5%)	β_2 (1%)	z	z (5%)
Part I	Odd	0.032	0.117	0.246	2.95	2.41	2.25	0.024	0.143
	Even	0.019			2.72				
Part II	Odd	0.009	0.117	0.246	2.55	2.41	2.25	0.104	0.143
	Even	0.059			2.27				
Part III	Odd	0.015	0.117	0.246	2.74	2.41	2.25	0.134	0.143
	Even	0.023			2.87				

VI. RELATIONSHIP BETWEEN THE 'RELIABILITY' AND THE
'SENSITIVITY' OF A TEST

It follows from equations (4) that

$$\rho = \frac{\gamma^2}{1 + \gamma^2}, \quad \dots\dots(28)$$

where ρ is the population value of the correlation coefficient between y_{1t} and y_{2t} or, what is usually called the reliability of the test. Table XI gives the values of γ corresponding to certain values of ρ . It is seen that for $\rho = 0.94$, σ is roughly $\sigma_e/4$, so that even for this large value of ρ we should conclude that the random errors of measurement may be very important in determining the actual score of the individual on the test.

Table XI

ρ	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
γ	0.00	0.33	0.50	0.66	0.82	1.00	1.22	1.53	2.00	3.00
ρ	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99	1.00
γ	3.18	3.39	3.64	3.96	4.36	4.90	5.69	7.00	9.95	∞

The use of γ , rather than ρ , as a measure of the reliability of the test has several advantages. I introduced the term 'sensitivity' of the test not only because the concept of sensitivity seemed to me to be of fundamental importance, but also in order to emphasize the desirability of using γ . First, from a mathematical point of view ρ and γ are equivalent, i.e. ρ is a function of γ ; but this is not the problem. The true value of γ , or ρ , is in any case unknown; our problem is to estimate the value of γ . It may be argued that we could estimate ρ , and then γ , but this introduces an unnecessary step and, what is more important, it is not the best method to use in estimating γ . The error of estimation of ρ would appear also in our estimate of γ , and the effect of this error on our estimate of γ is difficult to determine. The method of estimating γ given above is the simplest and the best that can be used.

Secondly, it is seen from Table XI that, for values of $\rho > 0.9$, a small change in the value of ρ corresponds to a large change in the value of γ . For example, a test for which $\rho = 0.99$ is more than twice as sensitive as a test for which $\rho = 0.95$. Thus it may happen that a test is actually much more sensitive than another but the difference in reliability might appear to be insignificant if our judgment is based on the values of the correlation coefficients.

Thirdly, the probability distributions of the statistics used in the above method are known, and are of a particularly simple form. The

probability distribution of r is known, of course, but it is difficult to evaluate the integral of the probability distribution in the cases for which $\rho \neq 0$ —i.e. the cases in which we should be interested if we wished to estimate ρ .

Finally, the methods outlined above could be generalized for values of $s > 2$ (i.e. in cases of more than two trials), but the correlation method is applicable only to the case $s = 2$. It is not necessary to use more than two trials, but it would probably be useful to do so, because we could then study the effect of other factors. For example, it would be interesting to consider the problem of possible changes in the variability of the students in repeated trials, i.e. to develop a test of the hypothesis that a common σ_e exists for all the trials. This problem is outside the scope of the present paper, but the same general method could be used in its solution.

VII. SENSITIVITY OF A 'WHOLE' TEST AND THAT OF A 'HALF' TEST

In the examples given above, the scores on Trial I and Trial II were obtained from the totals of the scores on the odd and even items of the test for each individual. Hence these are actually the scores on halves of the test, and the estimate of sensitivity obtained above is the estimate of the sensitivity of the half-test.

Denote by $\rho_{\frac{1}{2}}$ the correlation, in the population from which we are sampling, between the totals of the scores on the odd and even items of the test, and by ρ_w the correlation, in the population from which we are sampling, between the scores on two comparable forms of the same test. Further, denote by $\gamma_{\frac{1}{2}}$ the sensitivity of the half-test, and by γ_w the sensitivity of the whole test.

From formula (28) we see that

$$\rho_{\frac{1}{2}} = \frac{\gamma_{\frac{1}{2}}^2}{1 + \gamma_{\frac{1}{2}}^2} \quad \dots\dots(29)$$

and

$$\rho_w = \frac{\gamma_w^2}{1 + \gamma_w^2} \quad \dots\dots(30)$$

If the two forms of the test are comparable then it follows from the Spearman-Brown formula that

$$\rho_w = \frac{2\rho_{\frac{1}{2}}}{1 + \rho_{\frac{1}{2}}} \quad \dots\dots(31)$$

From (30) it follows that

$$1 + 2\gamma_w^2 = \frac{1 + \rho_w}{1 - \rho_w} \quad \dots\dots(32)$$

Substituting the value of ρ_w from (31) in (32) we have

$$1 + 2\gamma_w^2 = \frac{1 + 3\rho_{\frac{1}{2}}}{1 - \rho_{\frac{1}{2}}}. \quad \dots\dots(33)$$

Substituting the value of $\rho_{\frac{1}{2}}$ from (29) in (33) we have

$$1 + 2\gamma_w^2 = 1 + 4\gamma_{\frac{1}{2}}^2, \quad \dots\dots(34)$$

and we find that

$$\gamma_w = \sqrt{2} \gamma_{\frac{1}{2}}. \quad \dots\dots(35)$$

Hence to obtain the estimate of the sensitivity for the whole test from the results of experiments such as those considered above, it would be necessary to multiply the estimate of the sensitivity of the half-test by the factor $\sqrt{2}$. It must be remembered, however, that the above arguments are valid only if the two forms of the test are comparable—a condition which would be rarely satisfied in practice. For this reason the estimates of the sensitivity of the whole test calculated from the results obtained by the use of methods (1) or (2), and the estimate of the sensitivity for the half-test calculated from the results obtained by the use of method (3), will not satisfy equation (35) exactly.

Using equation (35), and the results from Table IX, we find the confidence intervals for γ_w which are shown in Table XII. It is seen that even these 'corrected' estimates of the sensitivity of the tests are rather small, and we should conclude, as before, that the random errors of measurement may be very important in determining the actual score of an individual on the tests.

Table XII

Test	Confidence interval for γ_w
Part I	$3.34 \leq \gamma_w \leq 5.13$
Part II	$2.98 \leq \gamma_w \leq 4.61$
Part III	$2.23 \leq \gamma_w \leq 3.54$

VIII. SUMMARY

The problem of the reliability of mental tests has been discussed using the theory of testing statistical hypotheses, and of estimation, devised and developed by Neyman and Pearson. Four separate problems have been considered:

- (1) to determine if there is a significant trial effect,
- (2) to determine if the mental test actually measures the capacity of the individuals tested,
- (3) to estimate the trial effect if it exists, and

(4) to obtain an estimate of the relative importance of the random errors of measurement, with respect to the true measurement of the capacity of the individual, in determining the score of the individual on the mental test.

A new concept, the sensitivity of a mental test, has been introduced as a measure of reliability in the solution of problems (2) and (4). The results for three mental tests, Comprehensive Examinations in Biological Sciences, have been analysed and discussed.

In conclusion I wish to thank Dr J. Neyman for the help which he has given me throughout the course of the work.

REFERENCES

- (1) NEYMAN, J. (1935). "Sur la vérification des hypothèses statistiques composées." *Bull. Soc. math. Fr.* LXIII.
- (2) — (1937). "Outline of a theory of estimation based on the classical theory of probability." *Philos. Trans.* (in the Press).
- (3) NEYMAN, J. & PEARSON, E. S. (1936). "Sufficient statistics and uniformly most powerful tests of statistical hypotheses." *Statist. Res. Mem.* Department of Applied Statistics, University College, London, I.
- (4) PEARSON, K. (1934). *Tables of the Incomplete Beta-Function.*
- (5) FISHER, R. A. (1932). *Statistical Methods for Research Workers.* Edinburgh: Oliver and Boyd.
- (6) GOODENOUGH, FLORENCE L. (1936). "A critical note on the use of the term 'reliability' in mental measurement." *J. Educ. Psychol.* XXVII, 173-8.
- (7) JOHNSON, P. O. & NEYMAN, J. (1936). "Tests of certain linear hypotheses and their application to some educational problems." *Statist. Res. Mem.* Department of Applied Statistics, University College, London, I.
- (8) THOULESS, R. H. (1935). "Test unreliability and function fluctuation." *Brit. J. Psychol.* XXVI, 325-43.
- (9) JACKSON, R. W. B. (1939). "Some problems of testing statistical hypotheses and estimation relating to the question of the relative accuracy of measurements." *Statist. Res. Mem.*, III.

(Manuscript received 25 August 1937)

THE INFLUENCE OF MULTIVARIATE SELECTION ON THE FACTORIAL ANALYSIS OF ABILITY

BY GODFREY H. THOMSON (SECTIONS I-V)
AND WALTER LEDERMANN (SECTIONS VI-VIII)
Moray House, University of Edinburgh

- I. *Introduction* (pp. 288-289).
- II. *Aitken's formula for multivariate selection* (pp. 289-290).
- III. *(Minimum) rank unchanged in the 'indirect' quadrants* (pp. 290-291).
- IV. *Creation and annihilation of group factors* (pp. 291-292).
- V. *The influence of selection on the factors themselves* (pp. 292-297).
- VI. *An algebraical proof of the theorem on multivariate selection* (pp. 297-304).
 - (a) *The induced correlations between factors* (pp. 297-299).
 - (b) *Analysis of the tests after selection* (pp. 299-304).
- VII. *The number of induced group factors* (pp. 304-305).
- VIII. *Application to the theory of univariate selection* (pp. 305-306).
References (p. 306).

I. INTRODUCTION

IN previous articles (1, 2) the influence of univariate selection on factorial analysis was considered, and it was shown that such selection leaves the rank of the 'reduced'* matrix of correlations unchanged although the loadings of all the factors are altered, i.e. one and the same test has different saturations with g , etc., in the population and the sample.

In view of what is to follow it is desirable to re-emphasize here what is stressed in the concluding sections of the previous paper, namely, that there is no real difference between population and sample in these considerations, and the relationships found work both ways. The word 'selection' suggests that all the individuals in the sample are contained

* By the reduced correlation matrix is meant in Thurstone's terminology the matrix with those quantities inserted in the principal diagonal which reduce its rank to a minimum. These quantities are called 'communalities', their defects from unity give the squares of the loadings of the specific factors, and the analysis for common factors is carried out on the reduced matrix. Throughout this paper I shall use the term 'minimum attainable rank' or for brevity 'minimum rank' to mean the lowest rank to which a matrix can be reduced by changing its leading diagonal.

among the individuals in the population. Even if that were so, a sample could be selected which had a larger standard deviation than the population. And in evolution, where in a certain sense one generation is a sample of the preceding, it may nevertheless outnumber its parents. It is better therefore to think of both sample and population as being samples from a hypothetical superpopulation of 'might-have-beens' (see (3), p. 183, quoted in (4), p. 188). It is clear then that there are as many analyses of a test as there are sample populations we can use it on, and that, mathematically at least, no one of these analyses has any claim to be considered the 'true' one. Univariante selection, however, does not alter the *number* of factors.

If now we turn to multivariate selection, where several variates are explicitly selected, we find that of course 'new factors' can be created, or old ones destroyed; but we shall see that these changes in the number of factors are only among the 'selected' variates, and that the remaining variates can still be analysed into the old number of factors, although with new loadings, and not in general identical in nature with the old factors.

II. AITKEN'S FORMULA FOR MULTIVARIATE SELECTION

Pearson's theory of multivariate selection has received considerable clarification from the formula into which it has been condensed by two recent papers of A. C. Aitken (5, 6). If $n = p + q$ tests are normalized their intercorrelations may be expressed in the $n \times n$ matrix

$$\begin{bmatrix} R_{pp} & R_{pq} \\ R_{qp} & R_{qq} \end{bmatrix}. \quad \text{.....(1)}$$

If now the variances and 'covariances'* of the first p tests are directly modified by 'selection' (in the wide sense) to new values V_{pp} then the rest of the matrix will be sympathetically modified to the form

$$\begin{bmatrix} V_{pp} & V_{pp} R_{pp}^{-1} R_{pq} \\ R_{qp} R_{pp}^{-1} V_{pp} & R_{qq} - R_{qp} (R_{pp}^{-1} - R_{pp}^{-1} V_{pp} R_{pp}^{-1}) R_{pq} \end{bmatrix}. \quad \text{.....(2)}$$

This is Aitken's formula: it is certainly applicable when the distributions are all normal, and it is believed to be applicable over a con-

* The word has presumably got too good a start to be recalled. But it already had a defined meaning in a closely allied branch of mathematics and its new use may lead to confusion. The new variances are referred to the new means if these have changed.

siderably wider range, though I understand that there is some doubt as to how wide this range is.

III. (MINIMUM) RANK UNCHANGED IN THE 'INDIRECT' QUADRANTS

As a preliminary example let us take the same matrix of correlations between five hierarchical tests as was used in (1), with an additional test added, namely,

$$\begin{bmatrix} 1.00 & 0.72 & 0.63 & 0.54 & 0.45 & 0.36 \\ 0.72 & 1.00 & 0.56 & 0.48 & 0.40 & 0.32 \\ 0.63 & 0.56 & 1.00 & 0.42 & 0.35 & 0.28 \\ 0.54 & 0.48 & 0.42 & 1.00 & 0.30 & 0.24 \\ 0.45 & 0.40 & 0.35 & 0.30 & 1.00 & 0.20 \\ 0.36 & 0.32 & 0.28 & 0.24 & 0.20 & 1.00 \end{bmatrix}, \quad \text{.....(3)}$$

and let us modify by intentional selection not only as there the first test, but the first two *and their covariance*, by changing the north-west quadrant of the matrix to

$$\begin{bmatrix} 0.36 & 0.30 \\ 0.30 & 0.36 \end{bmatrix}, \quad \text{.....(4)}$$

that is to say, we have selected a sample so that the individuals forming it are much less scattered both in test 1 and test 2 (variances 0.36 and 0.36 instead of 1.00) and also show a higher correlation between their scores in these two tests (0.83 instead of 0.72). Aitken's formula enables us then to write down what the variances and covariances of the remaining three tests become by reason of the correlations with the first two, namely,

$$\begin{bmatrix} 0.360 & 0.300 & 0.236 & 0.202 & 0.169 & 0.135 \\ 0.300 & 0.360 & 0.220 & 0.189 & 0.158 & 0.126 \\ 0.236 & 0.220 & 0.740 & 0.196 & 0.163 & 0.131 \\ 0.202 & 0.189 & 0.196 & 0.808 & 0.140 & 0.113 \\ 0.169 & 0.158 & 0.163 & 0.140 & 0.868 & 0.094 \\ 0.135 & 0.126 & 0.131 & 0.113 & 0.094 & 0.915 \end{bmatrix}, \quad \text{.....(5)}$$

which on being normalized gives the matrix of correlations

$$\begin{bmatrix}
 1.000 & 0.833 & 0.457 & 0.375 & 0.302 & 0.235 \\
 0.833 & 1.000 & 0.426 & 0.351 & 0.283 & 0.219 \\
 0.457 & 0.426 & 1.000 & 0.253 & 0.203 & 0.159 \\
 0.375 & 0.351 & 0.253 & 1.000 & 0.167 & 0.131 \\
 0.302 & 0.283 & 0.203 & 0.167 & 1.000 & 0.105 \\
 0.235 & 0.219 & 0.159 & 0.131 & 0.105 & 1.000
 \end{bmatrix}. \quad \dots\dots(6)$$

It is this last matrix which would be found by an experimenter who *began* with the sample. And examination shows that *it is still hierarchical in the three indirectly selected quadrants*. The only correlation which breaks the hierarchy is the 0.833 which we made intentionally of that size by selection. Similarly the matrix of variances and covariances is hierarchical except for the 0.300 of tests 1 and 2.

This discovery I was rapidly able to confirm by other examples of larger order and of higher rank. In every case, the minimum rank of the matrix, that is to say the rank to which it can be reduced by placing suitable 'communalities' in the leading diagonal, remained the same in the three quarters of the matrix *indirectly* selected, and was only changed in the directly selected quadrant.

IV. CREATION AND ANNIHILATION OF GROUP FACTORS

In our simple example we note that since the directly selected quadrant no longer conforms to the hierarchy, *a new group factor has been created by the selection*, and would be presumed to exist by an experimenter who began with what we are calling the 'sample'. The original correlation matrix could be analysed as follows:

$$\left. \begin{aligned}
 z_1 &= 0.9f + 0.436s_1, \\
 z_2 &= 0.8f + 0.600s_2, \\
 z_3 &= 0.7f + 0.714s_3, \\
 z_4 &= 0.6f + 0.800s_4, \\
 z_5 &= 0.5f + 0.866s_5, \\
 z_6 &= 0.4f + 0.917s_6,
 \end{aligned} \right\} \quad \dots\dots(7)$$

the new correlation matrix as follows:

$$\left. \begin{aligned} z_1 &= 0.824f_1 + 0.447f_1 + 0.348s_1, \\ z_2 &= 0.769f_1 + 0.447f_2 + 0.457s_2, \\ z_3 &= 0.555f_1 & + 0.832s_3, \\ z_4 &= 0.456f_1 & + 0.890s_4, \\ z_5 &= 0.366f_1 & + 0.931s_5, \\ z_6 &= 0.286f_1 & + 0.958s_6. \end{aligned} \right\} \dots\dots(8)$$

If the 'sample' is the whole of the present generation then the 'new' group factor f_2 is presumably as real as any other group factor. The population from which the sample is 'selected' may be inaccessible to us. It may be the preceding generation from whom our sample is descended. It may be the present generation extended by all those who died in infancy, or even by all those who were never born. It becomes abundantly clear on taking thought that *all* populations are samples.

The process of selection, and Aitken's formula, is reflexive, and our example can be worked backwards, illustrating the disappearance of a 'factor' through evolution or through actual human selection. The 'purification' of a hierarchical battery of tests has usually been carried out by evicting tests from the test sample. It could also be carried out by evicting persons from the population of subjects, a process which could be engineered so as to make the erring parts of the matrix conform to the hierarchy, without injuring the hierarchical perfection of the remainder.

V. THE INFLUENCE OF SELECTION ON THE FACTORS THEMSELVES

Shortly afterwards it occurred to me that a still wider use of Aitken's formula can be made, which throws yet more light on the nature of factors. The latter themselves are variates exactly as the tests are, and the analysis of the tests has given us the correlations of the factors with the tests and with each other (we are considering orthogonal factors but oblique factors could also be handled in the same way). If we write the factor correlations (both of group and specific factors) down exactly as though the factors were a continuation of the series of tests we obtain the matrix

$$\left[\begin{array}{c|cc} R_{pp} & R_{pa} & R_{pf} \\ \hline R_{ap} & R_{aa} & R_{af} \\ R_{fp} & R_{fa} & R_{ff} \end{array} \right], \dots\dots(9)$$

and the result of changing R_{pp} by selection to V_{pp} will be to give the matrix

$$\begin{bmatrix} V_{pp} & V_{pq} & V_{pf} \\ V_{qp} & V_{qq} & V_{qf} \\ V_{fp} & V_{fq} & V_{ff} \end{bmatrix}, \quad \dots\dots(10)$$

where

$$\begin{bmatrix} V_{pp} & V_{pq} \\ V_{qp} & V_{qq} \end{bmatrix}$$

is the same as the matrix (2), and

$$\left. \begin{aligned} V_{pf} &= V_{pp} R_{pp}^{-1} R_{pf}, \\ V_{qf} &= R_{qf} - R_{qp} R_{pp}^{-1} R_{pf} + R_{qp} R_{pp}^{-1} V_{pp} R_{pp}^{-1} R_{pf}, \\ V_{ff} &= R_{ff} - R_{fp} R_{pp}^{-1} R_{pf} + R_{fp} R_{pp}^{-1} V_{pp} R_{pp}^{-1} R_{pf}, \end{aligned} \right\} \quad \dots\dots(11)$$

and if the factors are originally orthogonal $R_{ff} = I$. Note that the actual calculations need not be divided up but can be done all in a piece.

The letter f above refers to *all* the factors. We can, however, if we wish divide the matrix up still further. If we henceforward restrict f to mean common or group factors, s to mean the specifics of the directly selected variates, and t to mean the specifics of the other variates, we can write the original matrix as

$$\begin{bmatrix} R_{pp} & R_{pq} & R_{pf} & R_{ps} & R_{pt} \\ R_{qp} & R_{qq} & R_{qf} & R_{qs} & R_{qt} \\ R_{fp} & R_{fq} & R_{ff} & R_{fs} & R_{ft} \\ R_{sp} & R_{sq} & R_{sf} & R_{ss} & R_{st} \\ R_{tp} & R_{tq} & R_{tf} & R_{ts} & R_{tt} \end{bmatrix}, \quad \dots\dots(12)$$

(where in the case of orthogonal factors R_{ff} , R_{ss} , and R_{tt} are all unit matrices, and R_{fs} , R_{ft} , R_{st} and their transposes zero matrices). The matrices after selection in the p variates will be represented by the same pattern with V instead of R (the suffixes unchanged), and it is convenient to mention here that Dr Walter Ledermann (who will use this arrangement in the later sections of this paper) has shown that Aitken's formula can be compressed into the statement, true for any number of cells in the above kind of dissection, that each submatrix (including V_{pp} itself) is given by the formula

$$\left. \begin{aligned} V_{\alpha\beta} &= R_{\alpha\beta} - R_{\alpha p} E_{pp} R_{p\beta} \quad (\alpha, \beta = p, q, f, \dots), \\ E_{pp} &= R_{pp}^{-1} - R_{pp}^{-1} V_{pp} R_{pp}^{-1}. \end{aligned} \right\} \quad \dots\dots(12a)$$

A numerical example will best illustrate the conclusions that can be drawn from this wider application of the selection formula. Let the original matrix of correlations of five tests be as shown in (13), together

with the correlations of these tests with their two common factors and five specifics, all the factors being orthogonal. Table (13) is in fact a complete representation of what Thurstone calls a 'structure'.

	<i>p</i>		<i>q</i>			<i>f</i>		<i>s</i>		<i>t</i>		
<i>p</i>	1.00	0.46	0.59	0.36	0.41	0.70	0.40	0.59	.	.	.	.
	0.45	1.00	0.36	0.26	0.23	0.60	0.10	.	0.79	.	.	.
<i>q</i>	0.59	0.36	1.00	0.32	0.45	0.50	0.60	.	.	0.62	.	.
	0.36	0.26	0.32	1.00	0.22	0.40	0.20	.	.	.	0.89	.
	0.41	0.23	0.45	0.22	1.00	0.30	0.50	.	.	.	.	0.81
<i>f</i>	0.70	0.60	0.50	0.40	0.30	1.00	.	.	.	.	.	.
	0.40	0.10	0.60	0.20	0.50	.	1.00	.	.	.	.	.
<i>s</i>	0.59	.	.	.	.	.	.	1.00	.	.	.	.
	.	0.79	.	.	.	.	.	.	1.00	.	.	.
<i>t</i>	.	.	0.62	.	.	.	.	.	.	1.00	.	.
	.	.	.	0.89	.	.	.	.	.	.	1.00	.
	.	.	.	.	0.81	.	.	.	.	.	.	1.00

.....(13)

The numbers in this table are supposed to be exact, except that the loadings of the specifics should be exactly the square roots of 0.35, 0.63, 0.39, 0.80, 0.66. In future tables only two places of decimals are printed, for clearness and economy of space, but the calculations were of course carried further.

Now let the variances and covariances of the first two tests be changed to the values shown in (14), upon which the rest of the matrix, *including the correlations of the factors*, will change sympathetically to the values there shown. The matrix of new correlations can at once be written down from (14), but we shall omit it to save space.

The first point of interest which arises from an inspection of this table (14) is that *the originally orthogonal factors are not now all orthogonal*. Those of them which are concerned in the composition of the directly selected variates have become correlated.

The second point of interest is that the minimum attainable rank of

the matrix remains unchanged within the square of the tests and their common factors, except for the corner square V_{pp} . Indeed the area of rank r (rank 2 in our example) has extended to include the cross-variances of the specifics of the selected tests with the other tests and the common factors, an area which formerly was blank, so that *the area of minimum attainable rank r forms a broad cross on the table*, as marked.

	p		q			f		s		t		
p	0.36	0.36	0.23	0.15	0.16	0.33	0.12	0.15	0.20	.	.	.
	0.36	0.64	0.27	0.19	0.17	0.43	0.09	0.05	0.48	.	.	.
q	0.23	0.27	0.80	0.20	0.31	0.28	0.45	-0.24	0.07	0.62	.	.
	0.15	0.19	0.20	0.93	0.14	0.26	0.12	-0.13	0.02	.	0.89	.
	0.16	0.17	0.31	0.14	0.90	0.15	0.39	-0.17	0.06	.	.	0.81
f	0.33	0.43	0.28	0.26	0.15	0.74	-0.15	-0.22	-0.01	.	.	.
	0.12	0.09	0.45	0.12	0.39	-0.15	0.88	-0.21	0.12	.	.	.
s	0.15	0.05	-0.24	-0.13	-0.17	-0.22	-0.21	0.65	0.26	.	.	.
	0.20	0.48	0.07	0.02	0.06	-0.01	0.12	0.26	0.59	.	.	.
t	.	.	0.62	.	.	.	.	.	.	1.00	.	.
	.	.	.	0.89	.	.	.	.	.	.	1.00	.
	.	.	.	.	0.81	.	.	.	.	.	.	1.00

.....(14)

Since the test square itself is still of minimum attainable rank r except for V_{pp} , it can be re-analysed into the same number of *orthogonal* common factors as before, plus new group factors d confined to the p directly selected tests, as is shown in (15) giving variances and covariances, and in (16) giving correlations.

There are seen in (15) and (16) two new orthogonal group factors e in place of the old and now oblique factors f ; a new group factor d running through the directly selected tests only; and new orthogonal specifics y to replace the specifics s which have become oblique. This is the manner in which an experimenter who began with the sample would analyse the

	<i>p</i>		<i>q</i>			<i>e</i>		<i>d</i>	<i>y</i>		<i>t</i>		
<i>p</i>	0.36	0.36	0.23	0.15	0.16	0.37	0.21	0.37	0.20	.	.	.	.
	0.36	0.64	0.27	0.19	0.17	0.49	0.20	0.37	.	0.48	.	.	.
<i>q</i>	0.23	0.27	0.80	0.20	0.31	0.33	0.55	.	.	.	0.62	.	.
	0.15	0.19	0.20	0.93	0.14	0.31	0.18	.	.	.	.	0.89	.
	0.16	0.17	0.31	0.14	0.90	0.17	0.46	.	.	.	.	.	0.81
<i>e</i>	0.37	0.49	0.33	0.31	0.17	1.00	.	.	.	.	.	.	.
	0.21	0.20	0.55	0.18	0.46	.	1.00	.	.	.	.	.	.
<i>d</i>	0.37	0.37	.	.	.	.	.	1.00	.	.	.	.	.
<i>y</i>	0.20	.	.	.	.	.	.	.	1.00	.	.	.	.
	.	0.48	.	.	.	.	.	.	.	1.00	.	.	.
<i>t</i>	.	.	0.62	.	.	.	.	.	.	.	1.00	.	.
	.	.	.	0.89	.	.	.	.	.	.	.	1.00	.
	.	.	.	.	0.81	.	.	.	.	.	.	.	1.00

.....(15)

tests (though he can ‘rotate’ his loadings) if he were wedded to orthogonal factors.

The concept of orthogonal factors is indeed even more fluid and dependent upon the sample than I realized in my former paper on univariate selection. After univariate selection the matrix can, it is true, be analysed into an orthogonal pattern of exactly the same form (position of zeros) as before, though the loadings have changed. But some at least of the factors have also changed, namely those concerned in the selected variate, for the previous factors are now correlated and the new ones are orthogonal. In the special case of *univariate* selection on a *hierarchical* matrix the change in nature can I think be confined to the specific of the directly selected test.

The discoveries described in the above were made partly empirically, partly by a loose use of various algebraic and geometric tools and a good deal of intuition. I therefore invited Dr Walter Ledermann to submit

	<i>p</i>		<i>q</i>			<i>e</i>		<i>d</i>	<i>y</i>		<i>t</i>		
<i>p</i>	1.00	0.75	0.44	0.26	0.28	0.62	0.34	0.62	0.34	.	.	.	.
	0.75	1.00	0.37	0.24	0.23	0.61	0.24	0.46	.	0.60	.	.	.
<i>q</i>	0.44	0.37	1.00	0.23	0.36	0.36	0.62	.	.	.	0.70	.	.
	0.26	0.24	0.23	1.00	0.15	0.32	0.19	.	.	.	.	0.93	.
	0.28	0.23	0.36	0.15	1.00	0.18	0.48	.	.	.	.	.	0.86
<i>e</i>	0.62	0.61	0.36	0.32	0.18	1.00	.	.	.	.	.	.	.
	0.34	0.24	0.62	0.19	0.48	.	1.00	.	.	.	.	.	.
<i>d</i>	0.62	0.46	.	.	.	.	.	1.00	.	.	.	.	.
<i>y</i>	0.34	.	.	.	.	.	.	.	1.00	.	.	.	.
	.	0.60	.	.	.	.	.	.	.	1.00	.	.	.
<i>t</i>	.	.	0.70	.	.	.	.	.	.	.	1.00	.	.
	.	.	.	0.93	.	.	.	.	.	.	.	1.00	.
	.	.	.	.	0.86	.	.	.	.	.	.	.	1.00

.....(16)

them to a rigorous examination and give exact proofs, and this, with some additional matter, he gives in the remaining sections of this paper, which are from his pen.

VI. AN ALGEBRAICAL PROOF OF THE THEOREM ON MULTIVARIATE SELECTION

(a) *The induced correlations between factors*

This section contains a formal proof of Prof. Godfrey H. Thomson's theorem on multivariate selection which has been set forth and illustrated by him in the first part of this paper. The theorem can be stated as follows (cf. p. 289):

If a team of $n=p+q$ tests has been resolved into r general and n specific factors which are normalized and mutually orthogonal, and if the variances and covariances of p of these tests are changed by selection, the mean (if also

changed) being again taken as the origin, then the battery can still be analysed into r generals and n specifics, but, in addition, a certain number l of group factors will appear which, however, run through the p selected tests only. Further these $r+n+l$ factors will again be uncorrelated and normalized.

Let $z = (z_p, z_q)$ be the vector of the $n = p + q$ tests, where z_p denotes the directly selected tests and where z_q refers to the remaining tests. We assume that the battery of tests has been analysed thus

$$\left. \begin{aligned} z_p &= R_{pf}f + R_{ps}s \\ z_q &= R_{qf}f + R_{qt}t, \end{aligned} \right\} \quad \dots\dots(17)$$

where f is a vector of r common factors, and s and t represent p and q specifics respectively, corresponding to the two groups of tests. The elements of the matrices R_{pf} , R_{ps} , R_{qf} and R_{qt} are the loadings of the general and specific factors and, in particular, R_{ps} and R_{qt} are diagonal matrices. The analysis (17) refers to a hypothetical superpopulation (see p. 289) which contains both the parent and the selected population as samples. Equations (17) therefore hold for the scores of each person under consideration, to whichever sample he may belong. We have suppressed a second subscript which would specify the individual person, and we shall rather regard each element of the column vectors $z_p, z_q, f, \dots$ as a row vector with N components corresponding to the number of persons in the sample. Moreover, we shall assume that the scale is chosen in such a way that the correlations are simply given by the scalar products of these row vectors. Thus we have

$$\left. \begin{aligned} R_{pf} &= z_p f', & R_{ps} &= z_p s', & R_{pt} &= z_p t' = 0, \\ R_{qf} &= z_q f', & R_{qs} &= z_q s' = 0, & R_{qt} &= z_q t', \end{aligned} \right\} \quad \dots\dots(18)$$

since z_p is independent of t , and z_q is independent of s , and remembering that the factors themselves are uncorrelated and normalized,

$$\begin{aligned} R_{ff} &= ff' = I, & R_{ss} &= ss' = I, & R_{tt} &= tt' = I, \\ R_{fs} &= fs' = 0, & R_{ft} &= ft' = 0, & R_{st} &= st' = 0. \end{aligned}$$

After selection, the variances and covariances are

$$\begin{aligned} V_{\alpha\beta} &= R_{\alpha\beta} - R_{\alpha p}(R_{pp}^{-1} - R_{pp}^{-1}V_{pp}R_{pp}^{-1})R_{p\beta} = R_{\alpha\beta} - R_{\alpha p}E_{pp}R_{p\beta}, \\ &\dots\dots(12a) \end{aligned}$$

where α and β take the values p, q, f, s, t (see p. 298). In particular, by means of (18), we find that

$$\left. \begin{aligned} V_{st} &= 0, & V_{ft} &= 0, & V_{tt} &= I, \\ V_{fs} &= -R_{fp}(R_{pp}^{-1} - R_{pp}^{-1}V_{pp}R_{pp}^{-1})R_{ps} = -R_{fp}E_{pp}R_{ps}, \\ V_{ff} &= I - R_{fp}(R_{pp}^{-1} - R_{pp}^{-1}V_{pp}R_{pp}^{-1})R_{pf} = I - R_{fp}E_{pp}R_{pf}, \\ V_{pf} &= V_{pp}R_{pp}^{-1}R_{pf}. \end{aligned} \right\} \dots\dots(19)$$

(b) *Analysis of the tests after selection*

We shall use a circumflex accent to denote the scores (referred to the new means) of the persons in the selected population. From the general relations (17) it immediately follows that

$$\left. \begin{aligned} \hat{z}_p &= R_{pf}\hat{f} + R_{ps}\hat{s}, \\ \hat{z}_q &= R_{qf}\hat{f} + R_{qt}\hat{t}, \end{aligned} \right\} \dots\dots(20)$$

where the coefficients $R_{pf}, \dots$, etc., are unaltered. We shall, however, assume that the new variances and covariances are again equal to the scalar products of the score vectors. This implies only a change of units; in fact, we have to multiply (17) throughout by $\sqrt{N/N'}$, where N' is the number of persons in the selected population. Thus we have

$$\hat{z}_p\hat{z}_p' = V_{pp}, \quad \hat{z}_p\hat{f}' = V_{pf}, \dots, \text{etc.} \quad \dots\dots(21)$$

In the selected population, the factors $\hat{f}, \hat{s}, \hat{t}$ are no longer orthogonal, and our aim is now to replace them by new factors, which shall be uncorrelated and normalized. It will be seen that this can, in general, be achieved only if further dimensions are added to the common-factor-space.

First, we shall split up $\hat{s}$ into a part which lies in the original common-factor-space and a part perpendicular to that space, thus

$$\hat{s} = X\hat{f}' + \hat{u}, \quad \dots\dots(22)$$

$$\text{where} \quad \hat{u}\hat{f}' = 0. \quad \dots\dots(23)$$

On multiplying (22) by $\hat{f}'$ we find

$$\hat{s}\hat{f}' = X\hat{f}\hat{f}',$$

$$\text{or, by (21),} \quad V_{sf} = XV_{ff},$$

$$\text{whence} \quad X = V_{sf}V_{ff}^{-1}.$$

Substituting this in (22) we get

$$\hat{s} = V_{sf}V_{ff}^{-1}\hat{f} + \hat{u}. \quad \dots\dots(24)$$

In general, we shall have

$$V_{ff} \neq I \quad (\text{see (19)}),$$

i.e. the variates $\hat{f}$ will be correlated in the selected population. But the matrix V_{ff} is a 'Gramian' matrix (being a correlation matrix) and is therefore capable of being factorized, thus

$$V_{ff} = KK', \quad \dots\dots(25)$$

where K is a square matrix with real elements. There are various methods by which the matrix K can be computed, e.g. L. L. Thurstone's 'diagonal method' ((7), p. 78), but the solution is not unique, a right-hand orthogonal factor being arbitrary. This indeterminacy is, however, inherent in every factor problem. After deciding on a particular solution K of (25) we put

$$\hat{f} = K\hat{e}, \quad \dots\dots(26)$$

where $\hat{e}$ represents r new group factors which are now uncorrelated and normalized; for we have

$$\hat{e}\hat{e}' = K^{-1}\hat{f}\hat{f}'K^{-1'} = K^{-1}V_{ff}K^{-1'},$$

$$\text{whence, by (25),} \quad \hat{e}\hat{e}' = I. \quad \dots\dots(27)$$

Substituting (24) and (26) in (20) we get

$$\left. \begin{aligned} \hat{z}_p &= (R_{pf} + R_{ps}V_{sf}V_{ff}^{-1})K\hat{e} + R_{ps}\hat{u}_s \\ \hat{z}_q &= R_{qf}K\hat{e} + R_{qt}\hat{t}. \end{aligned} \right\} \dots\dots(28)$$

The coefficient of $\hat{e}$ in the first equation can be more conveniently written as follows

$$\left. \begin{aligned} L &= (R_{pf} + R_{ps}V_{sf}V_{ff}^{-1})K = R_{pf}K + R_{ps}V_{sf}(KK')^{-1}K \\ &= R_{pf}K + R_{ps}V_{sf}K^{-1'} \quad (\text{by (25)}), \\ L &= (R_{pf}KK' + R_{ps}V_{sf})K^{-1'} \\ &= [R_{pf}(I - R_{fp}E_{pp}R_{pf}) - R_{ps}R_{sp}E_{pp}R_{pf}]K^{-1'} \quad (\text{by (19)}). \end{aligned} \right\} \dots\dots(29)$$

Now, the correlation matrices $R_{\alpha\beta}$ are not independent of one another, in virtue of the linear equations (17). Multiplying the first of these equations by its transpose we obtain

$$z_p z_p' = (R_{pf}f + R_{ps}s)(f'R_{fp} + s'R_{sp}),$$

and since the factors are mutually orthogonal and standardized in the original population

$$R_{pp} = R_{pf}R_{fp} + R_{ps}^2, \quad \dots\dots(30)$$

remembering that $R_{ps} = R_{sp}$ is a diagonal matrix. Using (30) to eliminate the product $R_{pf}R_{fp}$ we can write (29) in the form

$$L = [R_{pf} - (R_{pp} - R_{ps}^2) E_{pp} R_{pf} - R_{ps}^2 E_{pp} R_{pf}] K^{-1'},$$

and on cancelling the last two terms

$$L = [R_{pf} - R_{pp} E_{pp} R_{pf}] K^{-1'}.$$

Finally, on substituting for E_{pp} from (12a), p. 298, we get

$$L = [R_{pf} - R_{pp}(R_{pp}^{-1} - R_{pp}^{-1} V_{pp} R_{pp}^{-1}) R_{pf}] K^{-1'},$$

which readily simplifies to

$$L = V_{pp} R_{pp}^{-1} R_{pf} K^{-1'}.$$

Substituting this in (28) and putting, for brevity,

$$\hat{w} = R_{ps} \hat{u}, \quad \dots\dots(31)$$

we obtain

$$\left. \begin{aligned} \hat{z}_p &= V_{pp} R_{pp}^{-1} R_{pf} K^{-1'} \hat{e} + \hat{w}, \\ \hat{z}_q &= R_{qf} K \hat{e} + R_{qt} \hat{t}. \end{aligned} \right\} \quad \dots\dots(32)$$

The test scores are now expressed in terms of three sets of factors, viz. $\hat{e}$, $\hat{t}$ and $\hat{w}$. Factors belonging to different sets are orthogonal, for we have

$$\hat{e} \hat{t}' = K^{-1} \hat{f} \hat{t}' = K^{-1} V_{ft} = 0, \quad \text{by (26) and (19),}$$

$$\hat{e} \hat{w}' = K^{-1} \hat{f} \hat{u}' R_{sp} = K^{-1} (\hat{u} \hat{f}')' R_{sp} = 0, \quad \text{by (23),}$$

$$\hat{t} \hat{w}' = \hat{t} \hat{u}' R_{sp} = \hat{t} (\hat{s} - V_{sf} V_{ff}^{-1} \hat{f})' R_{sp} = \hat{t} \hat{s}' - \hat{t} \hat{f}' V_{ff}^{-1} V_{fs} R_{sp} = 0,$$

$$\text{since, by (19),} \quad V_{st} = V_{ft} = 0.$$

Moreover, the various components of $\hat{e}$ and of $\hat{t}$ are within themselves uncorrelated and normalized, since by (27) and (19)

$$\hat{e} \hat{e}' = V_{ee} = I,$$

$$\hat{t} \hat{t}' = V_{tt} = I.$$

We have, therefore, reached our goal except for the p factors $\hat{w}$ which will, in general, not be orthogonal. However, we could orthogonalize them by a process similar to that which was used to transform $\hat{f}$ into $\hat{e}$, and we should then obtain p new orthogonal group factors. But in accordance with the principle of minimizing the number of group factors at the expense of specific factors we shall endeavour to resolve $\hat{w}$ into as small a number l of group factors as possible and into p specifics (see § VII below), thus

$$\hat{w} = A_{wd} \hat{d} + B_{wy} \hat{y}, \quad \dots\dots(33)$$

where $\hat{d}$ represents the l group factors and $\hat{y}$ the p specifics; the elements of the matrices A_{pd} and B_{py} are the corresponding loadings, the matrix B_{py} being a non-singular diagonal matrix.

Now, (33) constitutes Thurstone's factor problem for the p fictitious tests $\hat{w}$ and the required analysis can be carried out by the usual methods. The matrix of variances and covariances of $\hat{w}$ which determines the factorial analysis of $\hat{w}$ will be denoted by V_{ww} , and an explicit expression for it will be derived forthwith. It should be noted that $\hat{w}$ is not in standard measure, but the process which turns V_{ww} into the actual correlation matrix (i.e. dividing each row and column by the respective standard deviation) only slightly modifies Thurstone's method of factorizing a team of tests. The number of new group factors $\hat{d}$ is determined by the 'minimum rank' of V_{ww} and after substituting from (33) in (32) we see that the number of dimensions of the total factor space has now risen to

$$r+p+q+l=r+n+l,$$

i.e. l dimensions more than before selection. We can therefore place the additional l group factors at right angles to each of the old factors $\hat{f}$, $\hat{s}$, $\hat{t}$, or $\hat{e}$, $\hat{w}$, $\hat{t}$, since the latter are linear functions of the former factors (see (26), (31) and (24)). From (33) it then follows that the new specifics $\hat{y}$ are likewise orthogonal to $\hat{e}$ and $\hat{t}$. For after post-multiplying (33) by $\hat{e}'$ we get

$$\hat{w}\hat{e}' = A_{wd}\hat{d}\hat{e}' + B_{wy}\hat{y}\hat{e}',$$

whence

$$\hat{y}\hat{e}' = 0,$$

since $\hat{w}\hat{e}' = 0$ and $\hat{d}\hat{e}' = 0$ and since B_{wd} is non-singular. In the same way it can be shown that

$$\hat{y}\hat{t}' = 0.$$

Thus from (32) and (33) we get

$$\left. \begin{aligned} \hat{z}_p &= V_{pp} R_{pp}^{-1} R_{pf} K^{-1} \hat{e} + A_{wd} \hat{d} + B_{wd} \hat{y}, \\ \hat{z}_q &= R_{qf} K \hat{e} + R_{qt} \hat{t}, \end{aligned} \right\} \dots\dots (34)$$

where $\hat{e}$ and $\hat{d}$ are group factors, and $\hat{y}$ and $\hat{t}$ are specifics. All these factors are uncorrelated and normalized.

We shall now derive an expression for V_{ww} in terms of R_{pp} , V_{pp} and R_{ps} , the loadings of the specifics in the first p tests. It will be shown that

$$V_{ww} = (R_{ps}^{-2} + V_{pp}^{-1} - R_{pp}^{-1})^{-1}. \quad \dots\dots (35)$$

By (32) we have $\hat{w} = \hat{z}_p - V_{pp} R_{pp}^{-1} R_{pf} K^{-1} \hat{e}$.

Hence

$$V_{ww} = \hat{w}\hat{w}' = (\hat{z}_p - V_{pp} R_{pp}^{-1} R_{pf} K^{-1} \hat{e}) (\hat{z}_p - \hat{e}' K^{-1} R_{fp} R_{pp}^{-1} V_{pp}).$$

Now $\hat{z}_p \hat{e}' = \hat{z}_p \hat{f}' K^{-1'} = V_{pf} K^{-1'} = V_{pp} R_{pp}^{-1} R_{pf} K^{-1'}$

(by (26) and (19)). Again,

$$\hat{e} \hat{e}' = I \quad \text{and} \quad KK' = V_{ff} = I - R_{fp} E_{pp} R_{pf}.$$

After expanding the expression for V_{ww} and using these results we obtain, on making some obvious simplifications,

$$V_{ww} = V_{pp} - V_{pp} R_{pp}^{-1} R_{pf} (I - R_{fp} E_{pp} R_{pf})^{-1} R_{fp} R_{pp}^{-1} V_{pp}.$$

For brevity, we shall write this

$$V_{ww} = V_{pp} - V_{pp} R_{pp}^{-1} T_{pp} R_{pp}^{-1} V_{pp}, \quad \dots (36)$$

where

$$T_{pp} = R_{pf} (I - R_{fp} E_{pp} R_{pf})^{-1} R_{fp}.$$

Writing the second factor of the left member as an infinite geometric progression* we find

$$T_{pp} = R_{pf} \left\{ \sum_{i=0}^{\infty} (R_{fp} E_{pp} R_{pf})^i \right\} R_{fp}. \quad \dots (37)$$

We now use the fact (see (30), p. 300) that

$$R_{pf} R_{fp} = R_{pp} - R_{ps}^2 = D_{pp}, \quad \text{say.}$$

The terms on the right-hand side of (37) then become in turn (we are dropping suffixes where no mistake is possible):

$$i=0: \quad R_{pf} R_{fp} = 0,$$

$$i=1: \quad R_{pf} (R_{fp} E_{pp} R_{pf}) R_{fp} = DED, \dots,$$

$$i=2: \quad R_{pf} (R_{fp} E_{pp} R_{pf}) (R_{fp} E_{pp} R_{pf}) R_{fp} = DEDED, \dots, \text{etc.}$$

Hence

$$T_{pp} = D + DED + DEDED + \dots,$$

$$T_{pp} = [I + (DE) + (DE)^2 + \dots] D,$$

and again using the formula for a geometric progression

$$T_{pp} = (I - DE)^{-1} D.$$

Substituting this in (36) we obtain

$$V_{ww} = V_{pp} - V_{pp} R_{pp}^{-1} (I - DE)^{-1} D R_{pp}^{-1} V_{pp},$$

$$V_{ww} = V_{pp} R_{pp}^{-1} (I - DE)^{-1} [(I - DE) R_{pp} - D R_{pp}^{-1} V_{pp}].$$

By (12a) and (38) we find

$$I - DE = I - (R_{pp} - R_{ps}^2) (R_{pp}^{-1} - R_{pp}^{-1} V_{pp} R_{pp}^{-1});$$

hence, after expanding and simplifying,

$$I - DE = V_{pp} R_{pp}^{-1} + R_{ps}^2 R_{pp}^{-1} - R_{ps}^2 R_{pp}^{-1} V_{pp} R_{pp}^{-1},$$

* It can be shown that the conditions for convergence are fulfilled, i.e. that the modulus of each latent root of $R_{fp} E_{pp} R_{pf}$ is less than unity ((8), p. 98).

$$\text{and} \quad (I - DE) R_{pp} - DR_{pp}^{-1} V_{pp} = R_{ps}^2,$$

whence

$$V_{ww} = V_{pp} R_{pp}^{-1} (V_{pp} R_{pp}^{-1} + R_{ps}^2 R_{pp}^{-1} - R_{ps}^2 R_{pp}^{-1} V_{pp} R_{pp}^{-1}) R_{ps}^2, \\ V_{ww} = \{R_{ps}^{-2} (V_{pp} R_{pp}^{-1} + R_{ps}^2 R_{pp}^{-1} - R_{ps}^2 R_{pp}^{-1} V_{pp} R_{pp}^{-1}) R_{pp} V_{pp}^{-1}\}^{-1}.$$

After multiplying out the factors in the curled bracket we finally get

$$V_{ww} = (R_{ps}^{-2} + V_{pp}^{-1} - R_{pp}^{-1})^{-1}. \quad \dots\dots(38)$$

Thus formula (35) is established, and the theorem stated on p. 297 is proved. Our results can be summarized by saying that the new factorial matrix of the battery of tests is given by

$$\hat{F} = \begin{bmatrix} V_{pp} R_{pp}^{-1} R_{pf} K^{-1'} A_{pd} & B_{py} \\ R_{pf} K & R_{qt} \end{bmatrix}, \quad \dots\dots(39)$$

where B_{py} and R_{qt} are diagonal matrices containing the loadings of the specifics. The matrix K is defined in (14) and the matrices A_{pd} and B_{py} are obtained by factorizing the $p \times p$ correlation matrix V_{ww} (35) into a minimum number l of group factors and p specifics, thus

$$V_{ww} = [A_{wd} : B_{wd}] [A_{wd} : B_{wd}]' = A_{wd} (A_{wd})' + B_{wd}^2. \quad \dots\dots(40)$$

VII. THE NUMBER OF INDUCED GROUP FACTORS

As was shown in the preceding section, the question as to whether or not group factors will be created by selection, depends entirely on the nature of the matrix V_{ww} . Using a result which is due to Thurstone ((7), p. 76) we can generally say that the number l of induced group factors will, at any rate, satisfy the inequality

$$l \leq \left\{ \frac{(2p+1) - \sqrt{(8p+1)}}{2} \right\}, \quad \dots\dots(41)$$

where $\{x\}$ denotes the least integer equal to, or above x . For example, if p is equal to 2 or 3, at most one additional group factor will appear. On the other hand, in certain cases, l may remain well below the limit set by (41) and we shall now investigate *under what conditions the number l may be zero*. This means that the minimum attainable rank of V_{ww} must be nil; i.e. that V_{ww} must be a diagonal matrix. But if V_{ww} is a diagonal matrix, so is

$$V_{ww}^{-1} = R_{ps}^{-2} + V_{pp}^{-1} - R_{pp}^{-1} \quad (\text{by (35)}).$$

Again, since R_{ps}^{-2} is a diagonal matrix, it follows that

$$V_{pp}^{-1} - R_{pp}^{-1}$$

must be a diagonal matrix. We have therefore proved the criterion:

Selection in p variables will not give rise to additional group factors, if and only if the matrices R_{pp}^{-1} and V_{pp}^{-1} are identical, except for the diagonal cells.

VIII. APPLICATION TO THE THEORY OF UNIVARIATE SELECTION

The results that were arrived at in the two articles on univariate selection (1,2) can, of course, be deduced as special cases from the above theory of multivariate selection ($p=1$).

As an example we shall here obtain the value of the communality h_1^2 of the directly selected variable. In the papers referred to it was found (p. 455, formula (4)) that

$$h_1^2 = \frac{H_1^2(1 - q_1^2)}{1 - H_1^2 q_1^2}, \quad \dots\dots(42)$$

where H_1^2 is the communality of the first test before selection and

$$p_1^2 = 1 - q_1^2 \quad \dots\dots(43)$$

is the new variance.

In the notation of the present article we have

$$R_{pp} = [1], \quad R_{ps} = [u_1], \quad V_{pp} = [p_1^2], \quad B_{wy} = [b_1], \quad \dots\dots(44)$$

all these matrices being simply scalar quantities. Since u_1 is the loading of the specific in the first test, the communality (before selection) is given by

$$H_1^2 = 1 - u_1^2. \quad \dots\dots(45)$$

Now from the first equation (34) we see that b_1 is the loading of the specific after selection, whence, on standardizing, we get

$$h_1^2 = \frac{p_1^2 - b_1^2}{p_1^2}.$$

Putting $p=1$ in the relation (41) we infer that $l=0$; i.e. *no new group factor appears in univariate selection*, and (40) simplifies to

$$V_{ww} = B_{wy}^2,$$

or, on using (35) and (44), to

$$(u_1^{-2} + p_1^{-2} - 1)^{-1} = b_1^2.$$

On substituting for b_1^2 we get

$$h_1^2 = 1 - p_1^2(u_1^{-2} + p_1^{-2} - 1)^{-1},$$

which, in virtue of (43) and (45), is readily seen to be equivalent to

$$h_1^2 = \frac{H_1^2(1 - q_1^2)}{1 - H_1^2 q_1^2}.$$

This proves (42). The communalities of the other variates and the new values of the correlation coefficients could be deduced in a similar manner.

REFERENCES

- (1) THOMSON, GODFREY H. (1938). *This Journal*, xxviii, 451-9.
- (2) LEDERMANN, WALTER (1938). *This Journal*, xxix, 69-73.
- (3) THOMSON, GODFREY H. (1920). *Psychol. Rev.* xxvii, 173-90.
- (4) BROWN, WM. & THOMSON, GODFREY H. (1921 and 1925). *The Essentials of Mental Measurement*. Camb. Univ. Press.
- (5) AITKEN, A. C. (1934). *Proc. Edinb. Math. Soc.* (2) iv, 106-10.
- (6) — (1936). *Proc. Edinb. Math. Soc.* (2) v, 37-40.
- (7) THURSTONE, L. L. (1935). *The Vectors of Mind*. Chicago.
- (8) MACDUFFEE, C. C. (1933). *The Theory of Matrices*. Berlin.

(*Manuscript received 3 November 1937*)

EDITOR'S NOTE

Correction: For BLACKIE, this *Journal*, xxix, p. 187,
read T. P. BLACK and JOHN MACKIE

PUBLICATIONS RECENTLY RECEIVED

A Source Book of Gestalt Psychology. Prepared by WILLIS D. ELLIS.
With an Introduction by Prof. K. KOFFKA. London: Kegan
Paul, Trench, Trubner and Co., Ltd. 1938. Pp. xiv + 403. 21s. net.

Dr Ellis offers his *Source Book* in the hope that it may supplement Prof. Koffka's *Principles* in effecting "the necessary transition from the German psychology of 1912 to the present day". He has prepared some thirty-five abridged transcriptions of studies originally published in Germany between 1915 and 1929 by the major prophets and protagonists of the *Gestalttheorie*. From this book, Dr Ellis suggests in the Preface, the reader may hope to gain "some knowledge of what underlies the *Gestalt* theory both in fundamental findings and, *above all*, in the manner of thinking, of addressing oneself to problems, and in the utilisation of results".

The selections comprising the *Source Book* are divided into three principal groups. The first section (General Problems) contains abstracts of three articles of general theoretical interest, of which two are by Wertheimer and one by Köhler, together with selected passages from the somewhat esoteric and hitherto untranslated *Physische Gestalten*. The second section (Special Problems) is primarily devoted to experimental applications of the *Gestalt* principle. Needless to relate, the most favoured topic is visual perception, and no fewer than nineteen out of the twenty-one strictly experimental papers deal with investigation in this field. Wertheimer is represented here by his pioneer exposition of the conditions of bi-dimensional organization, Fuchs by his interesting studies of transparency and assimilation, Gelb by a condensation of his well-known article on colour constancy, and Benary by his ingenious and important study of the influence of form on brightness contrast. The field of stroboscopic movement is fittingly honoured by the inclusion of four selections, providing serviceable summaries of the explorations of Ternus, Hartmann, Lindemann and Duncker. In the domain of animal psychology, we have abstracts of Köhler's original experiments on relational choice and of Hertz's work upon configurational perception in insects and birds. Outside the field of perception, however, the experimental material is admittedly scanty. Indications of the applicability of certain *Gestalt* conceptions to memory are nevertheless suggested by the inclusion of selections from the studies of Zeigarnik, Gottschaldt and Wulf. Dr Ellis must also be congratulated upon his wisdom in including five selections from the *Psychologische Analyse hirnpathologischer Fälle*. Much of this work of Gelb, Goldstein and Fuchs is obviously of first-rate importance, and it has hitherto been accorded far too little recognition by the ordinary laboratory psychologist. One may hazard the opinion that Fuch's observations upon completion phenomena in hemianopsia furnish by far the most convincing illustrations of configurational tendencies in visual organization.

A few primarily theoretical articles are also included in the second section. Of these, the most interesting are perhaps the two papers of Wertheimer on the thought processes. The final group of selections is exclusively composed of replies to the criticisms of contemporary workers. Koffka politely ferrets out inconsistencies in Benussi's Production Theory, whilst Köhler, in grimmer mood, parries the onslaughts of Rignano and G. E. Müller.

Dr Ellis deserves warm congratulation upon the competence with which he has carried through an ambitious and exacting project. Yet the outcome of his labour is open to serious criticism. Might not the compiler of this volume have performed a more valuable service had his initial aspirations been altogether less exalted? The *Source Book* irresistibly suggests a kind of family album of the *Gestalt* group, and it is tempting to picture those who have lived and worked in its tradition turning the pages slowly and sentimentally, whilst images of the Berlin *Psychologisches Institut*, before the era of *Gleichschaltung*, rise sadly and lovingly to mind. For the research

worker to-day, a rather more extensive collection of rather less abbreviated transcriptions would have possessed unparalleled value, and the omission of some, at least, of the theoretical papers would have made possible the inclusion of important recent work, such as that of Metzger and von Restorff. Alternatively, a volume of unabridged translations of the early and relatively inaccessible theoretical papers of Wertheimer, Köhler and Koffka would have received an enthusiastic welcome from those whose principal interests lie in the development of psychological theory. Taking the book as it is, however, the choice of material cannot altogether escape criticism. For instance, Hornbostel's whimsy, *The Unity of the Senses*, has long been available in a complete English translation, and it is difficult to believe that there can be reasons other than sentimental responsible for its inclusion in the present volume. Von Frey's essay on tactual sensibility is exceedingly slight, and one cannot avoid the suspicion that its presence here depends largely on the sympathy shown by this distinguished physiologist towards certain of the notions embodied in the *Gestalt* doctrine. Many will deplore the altogether inadequate representation of Koffka; his early work on perceived motion would surely appear more worthy of transcription than the rather questionable researches on the γ movement included in the present volume.

On the other hand, however, all serious students will welcome the opportunity of reading in English so much of the work of Wertheimer. Perhaps this book will emphasize more than any other the enormous debt which psychology, and not merely *Gestalt* Psychology, owes to this highly original and significant thinker. Dr Ellis deserves our sincere appreciation, not only for his translations of Wertheimer, but also for making available (particularly to those whose knowledge of German is severely taxed by the length and complexity of the average *Forschung* article) such serviceable abstracts of so many important experimental studies. His translations are uniformly competent; in one selection only (Lewin) is criticism likely to focus directly upon the translation.

The final section of the *Source Book* will probably alienate a not inconsiderable number of readers. Surely the publication of retorts to criticisms without the merest abstract of the original argument smacks strongly of the totalitarian attitude and is highly undesirable.

Primary Mental Abilities. By L. L. THURSTONE. University of Chicago Press; Cambridge University Press. 1938. Pp. x+121. 9s. net.

This is the first of a series of Psychometric Monographs issued by the Psychometric Society. The author of this monograph described, in *The Vectors of Mind* (1935), a method of factorial analysis which he has subsequently applied to a large body of experimental data, and has presented the results in this book. Without familiarity with the earlier work the reader would find difficulty in following certain sections in the present monograph, but if he is willing to accept without question the validity of the method of factorial analysis used, he will find the experiment and results obtained described in full, clear detail.

The subjects of the experiment were mostly university students, numbering 240 in all, with a modal age of 18; and the fact that their average intelligence (as measured by the psychological examination of the American Council on Education) was exceptionally high shows them to form a highly selected group. They were given a battery of fifty-seven tests, "assembled so as to represent a fairly wide range of the mental activities that are typical in current psychological tests, with special emphasis on those tests which are used as measures of general intelligence". The tests were tentatively grouped under the headings: Abstraction, Verbal Tests, Space, Form, Number, Numerical Reasoning, Verbal Reasoning, Space Reasoning, Rote Learning and an Unclassified group which included tests of Rhythm, Sound Grouping, Vocabulary, etc. These tests took fifteen hours to give and were spread over five days. The interest and attention of the subjects over this long period were ensured by using subjects who volunteered for the purpose and by promising them each an individual

vocational report on the test findings. A description and sample of each test is given in the monograph, which is very useful to the reader who wishes to examine critically the factors isolated in the subsequent analysis.

The first step in the factorial analysis of the data obtained from these tests was the computation of the 1596 coefficients of correlation between the test scores. This huge task was substantially decreased by calculating tetrachoric coefficients instead of the more usual product-moment coefficients. In justification of this the author states: "The most complete procedure would seem to be to normalize each of the distributions of raw scores and then to compute the product-moment coefficients. But this assumption that the given distributions of raw scores require normalizing before entering the correlation tables is also just the assumption that is implied by the tetrachoric coefficient. Hence, in using tetrachoric coefficients, we are estimating the product-moment coefficient for the normalized distributions of scores" (p. 58). Doubts may arise in the reader's mind about the validity of this substitution;¹ but this review is not the place for attempting a criticism of the technique of factorial analysis and the assumptions underlying it, so that beyond a caveat to the reader that he should satisfy himself that the methods used are legitimate, no further remarks about the statistical procedures used will be made.

The remainder of the analysis is completed according to the method explained in the author's earlier work and the following factors are isolated:

(1) *The S factor.* This factor is visual or spatial in character and appears to be a common element in thirteen of the tests used.

(2) *The P factor.* This is a perceptual factor of a particular type, i.e. the perceptual function here seems to be a facility in perceiving detail that is imbedded in irrelevant material.

(3) *The N factor.* A numerical factor similar to that identified by Spearman and Kelley.

(4) *The V factor.* This is a factor which can be called "verbal relations", with the reservation that still another verbal factor involves verbal material in a psychologically different manner.

(5) *The W factor.* The other verbal factor, which seems to have as its principal characteristic a fluency in dealing with words, and which is separate from the *V* factor, which is concerned with ideas and meanings.

(6) *The M factor.* The factor of memory.

(7) *The I factor.* An inductive factor, characteristic of tests which demand that the subject shall find a rule or principle for each item in the test.

There remain two factors interpretation of which is tentative and not so clear as that of the preceding seven. These are described by the author as:

(8) *The R factor.* The common characteristic of the tests which exhibited this factor was that the successful completion of the tests involved some form of restriction in the solution.

(9) *The D factor.* A deductive factor entering into four tests only.

The author concludes by a few illustrations of how the 'mental profiles' of certain subjects of the investigation may be related to the vocational wishes of those subjects.

The monograph shows a sincere and painstaking attempt to exemplify the value of the author's method of analysing the factors involved in test performance, and the results are interesting and suggestive to psychologists. It must be borne in mind, however, that the isolation and naming of a factor by any method of factorial analysis does nothing further than prove the existence of this factor in the particular set of test scores under survey, and does not necessarily mean that there is an exact psychological counterpart in the subjects who originally performed the tests. The performance of the same tests by other groups of subjects might quite conceivably yield different results. In this case in particular, a highly selected group of subjects was tested, and

¹ Cf. *An Introduction to the Theory of Statistics*, by Yule and Kendall, 11th ed., p. 253: "The product-sum correlation and the tetrachoric correlation are two entirely different measures with quite different properties. The one is in no sense an approximation to the other, and the two may differ largely."

it would be of great interest to see if the results were repeated, and the same factors isolated, if an unselected group of subjects (among whom the test intercorrelations would almost certainly be much larger) were given the same tests. Nor does the isolation and naming of a factor prove that a mental ability corresponding to it is possessed in fixed quantities by the subjects tested (and, arguing from them, by mankind in general), so that we can say of a certain person that he possesses abilities *A, B, C, D, ...* to certain specified degrees. We are far yet from being able to do this and it may even be argued that envisaging such a procedure is undesirable, since it is suggestive of a retrogressive step in the history of psychology. E. G. C.

Statistical Tables for Biological, Agricultural and Medical Research. By R. A. FISHER and F. YATES. London: Oliver and Boyd. 1938. Pp. 90. 12s. 6d. net.

This is a useful set of tables which should be of service to psychologists. It includes the well-known tables of Fisher's *Statistical Methods for Research Workers* (normal deviates, *t* for small samples, χ^2 , etc.), usefully extended by the inclusion of a new column showing values for $P=0.001$. The tables for use in the analysis of variance have been much extended by the inclusion of more values of *P* and also by the provision of a table from which *P* can be calculated direct from variance ratios without taking logarithms. Of particular value to psychologists should be Table XX for replacing ordinals in ranked data by standard scores. There are also tables of factorials up to 300 and of logs, squares, etc. Although these latter are easily available in less expensive tables, it considerably enhances ease of working to have them, as here, conveniently spaced on a large page.

Philippine Pagans: The Autobiographies of Three Ifugaos. By R. F. BARTON. London: Geo. Routledge and Sons, Ltd. 1938. Pp. xvii + 271. 15s. net.

This book inevitably challenges comparison with another recent set of autobiographies of native people, Marjory Perham's *Ten Africans*. From such a comparison it emerges badly. It is said that the three selected Ifugao natives, two men and one woman, were "allowed to tell their stories with a minimum of questioning", having been merely "asked to tell the things in their lives that they considered most important". But the author's interpolated and bracketed remarks seem to make it fairly clear that, wittingly or unwittingly, they were being pushed all the time in the direction of a rather detailed account of their sexual adventures. The book does not, in fact, give the impression of presenting any full and rounded account of life in any community, and although the biographies presented are undeniably lively in various ways, they have the appearance of a good deal of witting or unwitting influence of selection. On the other hand, the method, very briefly indicated and illustrated towards the end of the book, of asking native informants how they would solve certain social problems given them by the enquirer, is of considerable interest and might be extended and exploited to great advantage in other regions.

Reaction to Conquest: Effects of Contact with Europeans on the Pondo of South Africa. By MONICA HUNTER. London: International Institute of African Languages and Cultures (Oxford University Press). 1936. Pp. xx + 582. 30s. net.

Dr Monica Hunter has written an extremely comprehensive, absorbing and valuable study, of the utmost importance for the social psychologist who is looking for actual evidence concerning the problems of the contact of social groups. Both by

training and by sympathy she has a very thorough knowledge of the Pondo people, and she writes with evident authority and sincerity. She deals with Pondoland itself, with a representative urban community and with Bantu on European farms. She passes in detailed review most of the elements of their culture which have been definitely affected by contact with Europeans. She shows how different may be the effects of such contact as between individual members of either side concerned and as between more or less organized and official groups. She is throughout remarkably sane, sensible and balanced, and she manages to be all these without any sacrifice of vitality or interest. General Smuts writes an introduction to the book in which he says: "To me, who have been in contact with the native mind all my life, the book is a fascinating account of a native tribe, full of curious facts and sidelights, not only on the native mind but on human nature as a whole." The comment is justified. The volume is not merely one more monograph on a native people; it is a notable contribution to an understanding of some of the most important problems in current sociology and social psychology.

Personality: A Psychological Interpretation. By GORDON ALLPORT.
London: Constable and Co. 1938. Pp. xiv + 588. 16s. net.

This is the London edition of a book already reviewed in *this Journal* (xxviii, 1938, 472). It differs in no respect from the American publication, but the work is a very important contribution to general and to experimental psychology, and many English readers will be glad to know that it is now readily accessible.

Mental Conflicts and Personality. By MANDEL SHERMAN. London:
Longmans and Co. 1938. Pp. viii + 319. 12s. 6d. net.

It is urged that all major and most minor difficulties of personality are due to mental conflicts. At the same time it is an error to conclude from this that all conflicts are bad, in the sense that they must lead to diminished efficiency. Many are inevitable, and the real problems arise, not from the conflicts themselves, but from the way they are faced and the solutions that may be attempted. Dr Sherman deals first with the nature and genesis of conflicts, and then proceeds to illustrate and discuss the common types of conflict as they are to be found in contemporary individual and social behaviour. He considers culture conflicts, conflicts of inferiority and insecurity, sex conflicts, conflicts and neuroses, and conflicts and antisocial behaviour. As is inevitable in a short book covering a great amount of ground, a good many special topics hardly receive the attention they deserve. As a sensible introductory study, however, the book has much for which it can be recommended. It cannot be said to rise to any great heights of originality, but it is clear, objective in outlook, the case material is new and mostly very well selected, and the book as a whole will leave the intelligent student with a very good picture of the fields of study with which it is concerned.

A Pediatrician in Search of Mental Hygiene. By BRONSON CROTHERS.
London: Oxford University Press (Humphrey Milford). 1937.
Pp. xix + 271. 8s. 6d. net.

Dr Crothers describes himself as "a conventionally trained pediatrician who has kept up an interest in neurology". He is convinced that "it is impossible to deal wisely with children handicapped by disorders of the nervous system unless the educational and emotional elements of each situation are carefully considered. Both as a practitioner and as a teacher of medical students I have recognized with increasing clearness that pediatricians are, as a rule, quite unaware of many of the stimulating

and highly significant contributions of the psychiatrists, the psychologists, the teachers and the social workers in regard to children". The book is an attempt to argue the case for more psychology for the medical student, and a closer co-operation between the doctor and the psychologist and psychiatrist in the treatment of disease. It is practical in that it not only states the case, but shows how the co-operation regarded as necessary could be, and in some instances, has actually been achieved. The argument is sober and entirely free from exaggerated statement. Dr Crothers recognizes faults and failings on both sides and wisely does not refrain from stating these. It is an American book and the English reader will be particularly interested in the opportunity which it gives him to compare conditions in that country with those which obtain in his own. There appears to be less difference than is often supposed. Some of the details of teaching and hospital administration which are advocated would need alteration before they could well be adapted to the English situation, but the general argument is closely applicable here also, and it is greatly to be hoped that the aims which Dr Crothers states clearly and concisely may also direct the future development of pediatrics in England.

Crooked Personalities in Childhood and After. By R. B. CATTELL. London: Nisbet and Co., Ltd.; Cambridge University Press. 1938. Pp. xi + 215. 7s. 6d. net.

This somewhat unpromising title is given to what is, nevertheless, a very useful and readable introduction to the psychotherapy, mainly of children, but with reference also to adult adjustment. The author gives brief outlines of the theories of Freud, Jung and Adler, as they relate to clinical practice, and describes the Freudian 'mechanisms' in more detail. It is no doubt of value to attempt to relate the findings of theoretical and experimental psychologists to those of psychopathologists, and apply the results of the former to clinical diagnosis. But Dr Cattell's attempt is too superficial and uncritical to be helpful—in contrast with his treatment of his psychopathological material. The chapters on the Work of the Clinic and the Family Background form the best part of the book; and the relation of maladjustment to the life and values of society as a whole, especially with regard to the maladjustments of genius, suggests a useful corrective to the pursuit of 'normality' at any price.

Both the student and the layman will find this book most valuable in tackling the involvements of psychopathology, particularly as it emphasizes the importance of a thorough study of the individual circumstances in each case, rather than of adherence to any pre-conceived theory.

The Ego and the Mechanisms of Defence. By ANNA FREUD; translated by C. BAINES. International Psycho-analytical Library, No. 30. London: The Hogarth Press and the Institute of Psycho-analysis. 1937. Pp. x + 196. 9s. net.

There has been a gradual tendency in psycho-analytical circles to widen the scope of analysis, to include that of the ego and super-ego as well as that of the id, and thus to bring psycho-analytical practice more into line with that of non-psycho-analytical practising psychologists.

In this book the author deals with the non-sexual factors and discusses the defence mechanisms, both normal and pathological, employed by the ego in its struggle with the id and super-ego, and the consequent effects on behaviour. She shows that, from a psycho-analytical standpoint, it is not to be expected that the mere removal of restraining influences in early childhood will prevent the development of neurosis which is a consequence of the ego's fear of the strength of the instinctive forces in the id. She considers at length the many character changes which occur before and at

puberty; and discusses the asceticism and intellectualism of the adolescent. The exposition has that clarity one associates with Anna Freud, and the ideas—whether or no one agrees with the explanations—will be of interest, not only to analysts, but also to all who have to do with children.

Peter Kurten: A Study in Sadism. By GEORGE GODWIN. London: The Acorn Press. 1938. Pp. 58. 5s. net.

This book is to be taken as a foreword to the same author's translation of Prof. Karl Berg's full-dress study of the same topic which it is hoped may appear under the title of *The Sadist*. It is a brief, unvarnished study of this man who appears to have been a fairly reasonable, mild-mannered, even attractive person in everyday life, but who, in his night exploits of brutal murder and sexual outrage, was an uncontrolled monster of depravity. The book is, of course, for the specialist; but it contains a few points of general psychological interest. For example, why should Kurten have had the most detailed memory for every event of his most disgusting crimes? Why should he, even in ordinary life, have felt regret for his infidelities, and none for his brutalities? Is it true to maintain, as the author, following Prof. Berg, does maintain, that Kurten's crimes might have been committed by Baudelaire, by Verlaine, or even by Goethe, except that all these three had achieved a catharsis of words? Have words such power? Perhaps the fuller study may provide convincing answers.

They Always Come Back. By CICELY MCCALL. London: Methuen and Co. 1938. Pp. xxiv + 258. 7s. 6d. net.

It is no use writing an ordinary descriptive review of this book. The volume must be read. It is an account of first-hand experiences as an officer at Holloway and at the girls' Borstal institution at Aylesbury. Miss McCall is, however, no ordinary officer. She has psychological training and, what is much more to the point, she evidently possesses also genuine psychological insight. She tells a perfectly plain story of life in the prison and at the institution and tells it in an engaging personal manner, writing all the time, not about cases, but about people. There is no sentimentalizing and no high-flown moralizing. Undoubtedly the record is selected as any such record must be, but it seems utterly impossible that anybody can read this book—and it is greatly to be hoped that a vast number of people will read it—without a vivid realization that there is a great deal in the prison and Borstal systems of this country that is psychologically wrong. More than that, the reader will be left with clear ideas of *what* is psychologically wrong and of how a large amount of it could quite easily be set right. The book is certainly out for reform, but it is written throughout with exceeding good humour. Its influence ought to be wide and beneficent.

Play in the Infants' School. By E. R. BOYCE. London: Methuen and Co. 1938. Pp. xiii + 188. 5s. net.

In 1933 Miss E. R. Boyce was appointed head teacher of the Infants Department of the Raleigh School, Stepney, London, E. She found 300 children on the roll, with an age range from three to seven-and-a-half years. They were all drawn from a poverty stricken area, many of them under-nourished, physically and mentally retarded, restricted and none too happy. She decided that the school must be reorganized, the children given more freedom, their school tasks made more of the order of a joyful adventure than of mere duties, and their physical development as much considered as their conventional mental training. Here is the record of the experiment which she undertook. It carries the story up to 1936. It gives full detail of equipment, time-tables, methods, difficulties and successes. As Dr Susan Isaacs says in a brief

Foreword: "Miss Boyce has shirked none of the difficulties of her enterprise. She shares with the readers her doubts and hesitation, her gropings and her failures, as well as her successes. She assesses the achievements of the school in ordinary scholastic terms, but is able to do more than this: she shows how for these children the level of actual attainment is far less significant than the fostering of the will to learn and the belief that learning has some value for the personal life." The story is fascinating and is well-told. Now it is greatly to be hoped that at least some of these children will be followed out of their Infants' school into their Primary and Secondary school life, so that it may become clear to what extent the lessons of adjustment and enthusiasm which seem to have been very successfully learned in the early stages persist in the larger world. There certainly appears every reason to hope and believe that they may do so.

The Assistant Master Speaks. Twelve Essays by Assistant and House Masters at Public and Secondary Schools. 1938. London: Kegan Paul, Trench, Trubner and Co., Ltd. Pp. viii+296. 7s. 6d. net.

These essays were obtained by inviting contributions from all Assistant Masters in Secondary Schools. It is not stated who made the selection here given; and the twelve essays are of very unequal merit. Two good constructive contributions on the improvement of the curriculum in Secondary Schools are made by C. D'O. Gowan and G. F. Golding, of Eton College and St Olave's Grammar School, London, respectively; and there is an excellent essay on Discipline by a House Master at an English Public School. These are full of sound psychology and practical wisdom. The other essays in the main reflect the dissatisfaction of school masters and of the general public with the present system of secondary education for boys, and make certain suggestions for remedying this state of affairs. But with one or two exceptions (notably that of Donald Hughes, of the Leys School) they are lacking in conviction.

An Introduction to a Christian Psycho-Therapy. By J. A. C. MURRAY. Edinburgh: T. and T. Clark. 1938. Pp. xii+279. 7s. 6d. net.

The book is a sincere and well-written argument, influenced in the main by a somewhat eclectic attitude towards modern psycho-therapeutic practice and theory. It is less a statement of the case for grafting psycho-therapy on to Christian thought and practice as for treating psycho-therapy as a natural outgrowth of that part of the Christian religion which is directly concerned with the life of people in the actual world. It is bound to make certain assumptions, one of which, for instance, is the old one of a religious instinct, but if these are granted the treatment runs smoothly and clearly and is supported by many references to original Christian teaching. The author considers that the time is ripe for the training of specially selected candidates for ordination in academic, experimental and especially clinical psychology. There can be little doubt that, if Christian beliefs and psycho-therapeutic practice can be happily combined, in many instances the strength and practical success of both may be greatly increased. But whether it is theoretically necessary, or even possible, for psycho-therapy to have a fundamentally Christian basis is much more disputable. Meanwhile the book may have a good influence provided it does not set off inadequately equipped persons in the immediate practice of the techniques which it describes.

REVENUE ACCOUNT AND BALANCE SHEET AS AT 30 SEPTEMBER 1938

Revenue Account for the year ended 30 September 1938

£1363	2	6
£1449	8	2

[P.T.O.]

REVENUE ACCOUNT AND BALANCE SHEET (continued)

Balance Sheet as at 30 September 1938

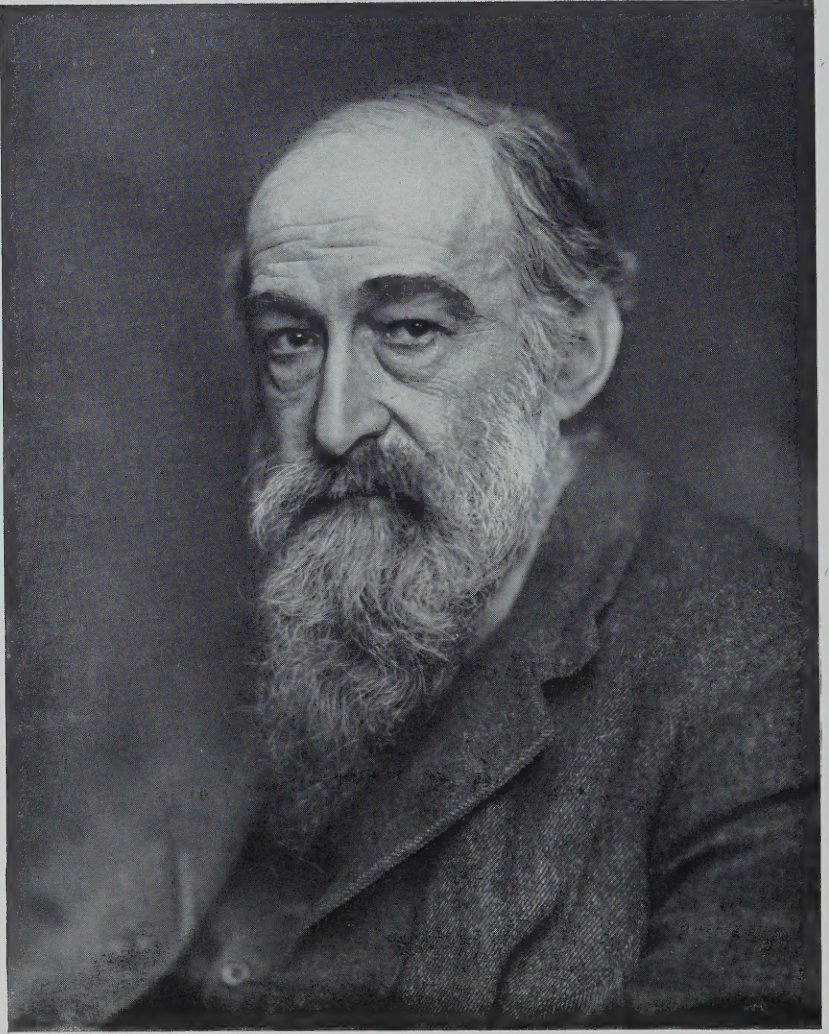
1937			1937			ASSETS						
£	s.	d.	£	s.	d.		£	s.	d.	£	s.	d.
1 1 0			238	8	6	Cash at Barclays Bank Ltd.	...	...	...	210	2	1
			1	3	0	Postage stamps in hand ...	...	...	...		16	6
122 0 0						<i>Investments at Cost:</i>						
						£300 5% Conversion Loan 1944-64...	...	...	318	9	9	
						£320. 14s. 8d. 3½% Inscribed War Loan ...	...	...	3425	7	9	
4765 14 1			3543	17	6	Note: Market value of above Securities at 30 September 1938, £3720				3743	17	6
						<i>Sundry Debtors</i>						
			1	9	3	Subscriptions in Arrear ...	...	...	...	129	15	0
			111	6	0	<i>Office Furniture and Fittings:</i>						
						Balance as per last Balance Sheet ...	...	...	177	15	0	
						Additions during year ...	...	...	13	10	9	
						Less Depreciation ...	...	...	191	5	9	
						Library:			19	2	9	
						Balance as per last Balance Sheet ...	...	...	583	0	0	
						Add Expenditure during year ...	...	...	57	5	7	
						Less Depreciation ...	...	...	640	5	7	
						<i>Monograph Publications:</i>			64	0	7	
						Balance as per last Balance Sheet ...	...	...	181	15	10	
						Less Amounts repaid during year ...	...	...	34	7	4	
						Shares in the British Journal of Educational Psychology, Ltd., at Cost ...	...	...				

I beg to report that I have examined the foregoing Accounts with the Books and vouchers of the British Psychological Society and that I have obtained all the information and explanations I have required. In my opinion the Balance Sheet is properly drawn up so as to exhibit a true and correct view of the state of the Society's affairs according to the best of my information and the explanations given to me and as shown by the Books of the Society. The Securities appearing under the heading Investments have been verified by production of the Bankers' Certificate.

T. BARRETT,
Incorporated Accountant

28 October 1938.

S. J. F. PHILPOTT,
Hon. Treasurer.



PROF. SAMUEL ALEXANDER, 1859-1938